

**Copyright**

**by**

**Kyle James Daun**

**2003**

**The Dissertation Committee for Kyle James Daun  
certifies that this is the approved version of the following dissertation:**

**Design Optimization of Radiant Enclosures**

**Committee:**

---

John R. Howell, Supervisor

---

Kenneth S. Ball

---

Ofodike A. Ezekoye

---

K. G. T. Hollands

---

David P. Morton

**Design Optimization of Radiant Enclosures**

by

**Kyle James Daun, B. Sc., M.A.Sc.**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

December, 2003

## Acknowledgements

I would foremost like to thank my advisor, Dr. John Howell, for his guidance and support throughout my doctorate, and to Dr. David Morton for introducing me to the subject of nonlinear programming. Throughout my academic career I have been fortunate to have exceptional mentors to guide both my professional and personal development; in addition to Drs. Howell and Morton, I am very grateful to Dr. Terry Hollands at the University of Waterloo, Drs. Ofodike Ezekoye, Janet Ellzey, and Philip Schmidt at The University of Texas at Austin, and Drs. Scott Ormiston and Hassan Soliman at the University of Manitoba. Thanks are also due to Ruth Schwab, Desiree Kane, and Dena Wagner, for keeping me registered and appointed in good standing despite my best efforts.

The graduate students in mechanical engineering made life at UT both interesting and enjoyable. This is particularly true of my friends from ETC 7.144, including Sean Bulla, John Dunn, Hakan Erturk, Francis Franca, Mirko Gamba, Charles Lan, Erik Wassen, Daijiao Wang, Dongmei Zhou, Alex Heltzel, Trey Pavy, Jason Porter, and Chang Rhee. (I'll always cherish the memories of Francis's Bach and Tea recitals, our hiking and canoeing expeditions, and Charles in Charge.)

This dissertation would not have been possible without the love, nurturing, and support I received from my family in Canada. Families take many forms, however; upon arriving in Texas, I could not have imagined the great people I have met who helped and supported me throughout my stay in Austin, and I'm sure they cannot imagine how important their friendship was to me. I am especially grateful to Eliana Schonberg, Jeff



Zuba, and David Alexander, as well as Caryn, Jami and Veena, Ted and Desiree, Sean and Nichole, Sherry, Clay, Clyde, Sharon, Rob, Greg, Lesslie, Michael, Jonathan, Amanda, Andy and Gretchen, and Ben and Laney.

## **Design Optimization of Radiant Enclosures**

Publication No. \_\_\_\_\_

Kyle James Daun, Ph. D.  
The University of Texas at Austin, 2003

Supervisor: John R. Howell

Design problems involving radiant enclosures are encountered in many different industrial applications. Examples include the design of annealing furnaces used in materials processing, ovens that bake food or cure coated surfaces, and rapid-thermal-processing chambers used to manufacture semiconductor wafers.

In each of these applications, the objective of the design problems is to find the enclosure geometry and heater settings that produce the desired temperature and heat flux distribution over the product. Traditionally, this has been done using a forward “trial-and-error” design methodology, which is a time-consuming process that results in a solution of limited quality. More recently, inverse design methodologies have been developed that require far less time than the forward methodology, and produce solutions

that better satisfy the desired conditions over the product surface. It is difficult to enforce design constraints, however, which often limits the usefulness and applicability of solutions obtained by this approach.

This dissertation describes several optimization methodologies that can be used to solve several types radiant enclosure design problems. In this approach, an objective function is first defined that quantifies the “goodness” of a particular design, in such a way that its minimum corresponds with the ideal design outcome. The objective function is dependent on a set of design parameters that control the enclosure configuration. Once this is done, the optimal set of design parameters is found by minimizing the objective function through nonlinear programming. Far less design time is required compared to the forward methodology, and the final solution is near-optimal. Furthermore, unlike the inverse methodology, it is possible to implement constraints by restricting the domain of the design parameters, which ensures that the solution can be easily implemented in a practical setting. Design methodologies are presented for design the heater settings and geometry of both diffuse-walled enclosures and enclosures containing surfaces with directionally dependent properties, and for solving the heater settings in problems involving transient and multimode heat transfer effects.

## Table of Contents

List of Tables		xi
List of Figures		xii
Nomenclature		xv
<b>Chapter 1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation	2
1.2	Literature Survey	11
1.2.1	Use of Inverse and Optimization Techniques in Experimental Settings	11
1.2.2	Use of Inverse Techniques to Solve Design Problems	13
1.2.3	Use of Optimization Techniques to Solve Design Problems	15
1.3	Dissertation Outline	18
<b>Chapter 2</b>	<b>Heater Setting Optimization for Radiant Enclosures Containing Diffuse-Gray Surfaces</b>	<b>20</b>
2.1	Introduction	21
2.2	Forward Design Methodology	25
2.3	Inverse Design Methodology	29
2.4	Optimization Design Methodology	35
2.4.1	Optimization Strategy	36
2.4.2	Calculation of Design Sensitivities	39
2.5	Demonstration of Inverse and Optimization Design Methodologies	41
2.5.1	Inverse Design Methodology Solution	43
2.5.2	Optimization Design Methodology Solution	48
2.5.3	Comparison of Inverse and Optimization Design Methodology Solutions	50
2.6	Conclusions	52
<b>Chapter 3</b>	<b>Geometric Optimization of Radiant Enclosures Containing Diffuse Surfaces</b>	<b>54</b>
3.1	Introduction	55
3.2	Optimization Strategy	59
3.3	Calculation of Design Sensitivities	61
3.4	Solution of the Governing Equations	66
3.5	Parametric Representation of Enclosure Surfaces through NURBS	71

	3.6	Implementation and Demonstration	76
	3.7	Conclusions	85
<b>Chapter 4</b>		<b>Geometric Optimization of Radiant Enclosures Containing Non-Diffuse Surfaces</b>	<b>87</b>
	4.1	Introduction	88
	4.2	Optimization Strategy	92
	4.3	Solution of the Governing Equations	97
	4.4	Implementation and Demonstration	102
	4.5	Conclusions	112
<b>Chapter 5</b>		<b>Optimization of Radiant Enclosures with Transient and Multimode Heat Transfer</b>	<b>113</b>
	5.1	Introduction	114
	5.2	Optimization Methodology for Radiant Enclosures involving Conduction, Convection, and Transient Heat Transfer	121
	5.2.1	Optimization Strategy	121
	5.2.2	Calculation of the Temperature Distribution and Sensitivities	128
	5.2.3	Implementation and Demonstration	136
	5.3	Optimization Methodology for Low-Thermal Inertia Furnaces	145
	5.3.1	Optimization Strategy	147
	5.3.2	Calculation of Net Radiant Heat Flux Distribution and Sensitivities	149
	5.3.3	Verification of Optimal Heater Settings	151
	5.3.4	Implementation and Demonstration	153
	5.4	Conclusions	163
<b>Chapter 6</b>		<b>Conclusions</b>	<b>165</b>
	6.1	Conclusions	166
	6.2	Recommendations for Future Work	174
	6.3.1	Application of MINLP and Metaheuristic Techniques	174
	6.3.2	Development of Optimization Methodologies for Enclosures Containing Participating Media	177
	6.3.3	Integration of a CFD Code into the Optimization Design Methodology	178
<b>Appendix A</b>		<b>Optimization through Nonlinear Programming</b>	<b>181</b>
<b>Appendix B</b>		<b>Solution of Inverse Radiant Enclosure Design Problems through Regularization</b>	<b>217</b>

Bibliography	253
Vita	260

## List of Tables

Table 2.1	Heater settings found with the inverse design methodology [ $W/m^2$ ]	46
Table 2.2	$L_2$ norms of the solution and residual vectors corresponding to solutions found using the inverse design methodology.	46
Table 2.3	Heater settings found with the optimization design methodology [ $W/m^2$ ].	48
Table 3.1	Grid refinement statistics for initial enclosure configurations shown in Figs. 3.11 (a) and 3.13 (a).	79
Table 5.1	Surface properties corresponding to the enclosure of Fig. 5.4.	137
Table 5.2	Design surface properties for the problem shown in Fig. 5.15.	154

## List of Figures

Figure 1.1	Example of a radiant enclosure design problem.	3
Figure 1.2	Boundary conditions for the forward design methodology.	5
Figure 1.3	Boundary conditions for the inverse design methodology.	6
Figure 1.4	Boundary conditions for the optimization design methodology.	8
Figure 1.5	Types of radiant enclosure design problems solved using inverse and optimization design methodologies in the literature: (a) inverse boundary condition estimation, (b) inverse source estimation, (c) transient problems, and (d) geometry design problems.	17
Figure 2.1	Parametric representation of the enclosure cross-section.	26
Figure 2.2	Discretization of the parametric domain.	27
Figure 2.3	Example of an inverse radiant enclosure design problem posed in its explicit form.	30
Figure 2.4	Example radiant enclosure design problem.	41
Figure 2.5	Grid refinement study for the design problem in Fig. 2.4, performed at initial configuration for optimization, $\Phi^0$ .	43
Figure 2.6	Ordered singular values found by decomposing $\mathbf{A}$ matrix.	45
Figure 2.7	Heat flux and emissive power distributions over the heater surface found using all the non-zero singular values ( $p = 449$ ).	47
Figure 2.8	Heat flux and emissive power distributions over the heater surface found using $p = 448$ singular values.	47
Figure 2.9	Heat flux and emissive power distributions over the heater surface found through the optimization design methodology.	49
Figure 2.10	Heat flux distributions over the design surface produced by the heater settings found using the optimization design methodology and the inverse design methodology with negative heater settings set equal to zero.	49
Figure 3.1	Example of a radiant enclosure inverse geometry design problem.	60
Figure 3.2	Geometric quantities used to form $k(u, u', \Phi)$ .	63
Figure 3.3	Calculation of $dF_{i-stripj}(\Phi)$ .	67
Figure 3.4	Example of a Bézier curve.	72
Figure 3.5	Basis functions corresponding to the Bézier curve in Fig.3.4.	72
Figure 3.6	Recursive formulation of cubic Bézier basis functions.	73
Figure 3.7	Example design problem.	77
Figure 3.8	Plot of $F(\Phi)$ over the feasible region.	78
Figure 3.9	Grid refinement studies for the design problem shown in Fig. 3.7.	79



Figure 3.10	Minimization path starting from $\Phi^{0,1} = \{0.8, 1.3\}^T$ and ending at $\Phi^{*,1} = \{1, 0.5896\}^T$ .	80
Figure 3.11	Enclosure geometries corresponding to (a) $\Phi^{0,1} = \{0.8, 1.3\}^T$ and (b) $\Phi^{*,1} = \{1, 0.5896\}^T$ .	81
Figure 3.12	Minimization path starting from $\Phi^{0,2} = \{0.8, 1.3\}^T$ and ending at $\Phi^{*,2} = \{0, 0.5\}^T$ .	82
Figure 3.13	Enclosure geometries corresponding to (a) $\Phi^{0,2} = \{0.5, 1\}^T$ and (b) $\Phi^{*,2} = \{0, 0.5\}^T$ .	82
Figure 3.14	Heat flux distributions corresponding to the initial and optimal enclosure configurations shown in Fig. 3.11 and Fig. 3.13.	83
Figure 4.1	Example radiant enclosure design problem, where some surfaces have directionally-dependent properties.	93
Figure 4.2	Random variables contained in $\xi$ control the location and direction of bundle emission from a surface element in a two-dimensional enclosure.	99
Figure 4.3	Radiant enclosure design problem involving two adiabatic, specularly-reflecting surfaces.	103
Figure 4.4	Grid refinement study for the design problem of Fig. 4.3.	104
Figure 4.5	Effect of the number of bundles on $\delta_l(\Phi^0)$ for the design problem of Fig. 4.3.	105
Figure 4.6	Minimization path for the design problem of Fig. 4.3.	106
Figure 4.7	Initial and final enclosure geometries for the design problem of Fig. 4.3.	107
Figure 4.8	Initial and final design surface heat flux distributions for the design problem of Fig. 4.3.	107
Figure 4.9	Radiant enclosure design problem involving multiple specular-diffuse surfaces.	110
Figure 4.10	Figure 4.10: Enclosure geometries corresponding with $\Phi^{0,1}$ and $\Phi^{*,1}$ .	110
Figure 4.11	Enclosure geometries corresponding with $\Phi^{0,2}$ and $\Phi^{*,2}$ .	111
Figure 4.12	Heat flux distributions over the design surface produced by the initial and local optimal enclosure geometries	111
Figure 5.1	Example of a transient radiant enclosure design problem: (a) radiant enclosure, and (b) desired set-point temperature history.	122
Figure 5.2	Relationship between $Q_{t-1 \rightarrow t}^{added}(\Phi)$ , $Q_{t-1 \rightarrow t}^{target}$ , $Q_{0 \rightarrow t}^{deficit}(\Phi)$ and the design surface temperature.	124
Figure 5.3	Infinitesimal wall element used to form governing temperature equation.	129

Figure 5.4	Discretization of the radiant enclosure into finite wall elements.	131
Figure 5.5	Example design problem involving conduction, convection, and transient heat transfer effects. (Heater numbers are shown in circles.)	137
Figure 5.6	Optimal heater settings for the problem shown in Fig. 5.5. (Heater numbers shown in circles.)	141
Figure 5.7	Variation of the design surface temperature from the set-point temperature throughout the process, for the problem shown in Fig. 5.5.	141
Figure 5.8	Temperature distribution over the design surface at different process times for the problem shown in Fig. 5.5.	142
Figure 5.9	Spatial grid refinement study at $\Phi^*$ , for the problem shown in Fig. 5.4, using $N_t = 500$ time steps.	143
Figure 5.10	Temporal grid refinement study at $\Phi^*$ , for the problem shown in Fig. 5.4, using $N = 240$ wall elements.	144
Figure 5.11	Example of a low-thermal inertia radiant enclosure: (a) enclosure boundary conditions and (b) relationship between radiosity, net radiant heat flux, and temperature increase.	148
Figure 5.12	Rapid thermal processing furnace design problem.	154
Figure 5.13	Optimal heater settings obtained using $N = 500$ wall elements, and $N_t = 100$ time steps, for the problem shown in Fig. 5.12.	157
Figure 5.14	Values of $F[\Phi^*(\tau_i)]$ corresponding to the heater settings shown in Fig. 5.13, and the number of time steps required to reach $\Phi^*(\tau_i)$ .	157
Figure 5.15	Variation of the design surface temperature from the set-point temperature throughout the process, for the problem shown in Fig. 5.12.	158
Figure 5.16	Non-dimensional temperature distribution over the design surface at different process times, for the problem shown in Fig. 5.12.	159
Figure 5.17	Refinement study on the number of surface elements used in the optimization and verification procedures.	161
Figure 5.18	Refinement study on the number of time steps used in the verification procedure.	162
Figure 6.1	Two possible heater surface configurations for a radiant enclosure design problem where the number, size, and location of radiant heaters are design variables.	175
Figure 6.2	Example design problem involving a radiant enclosure containing a participating medium. Design parameters control the heat output and location of multiple burners located within a participating medium.	178

## Nomenclature

$a$	Upper bound of the parametric region
$A_{DS}$	Design surface area
$\mathbf{A}$	Coefficient matrix
$b$	Lower bound of the parametric region
$b(u, \Phi)$	See Eq. (2.3) and (2.4), or Eq. (5.21)
$\mathbf{b}$	Right-hand side vector
$\mathbf{b}'$	Right-hand side vector for first-order sensitivity
$\mathbf{b}''$	Right-hand side vector for second-order sensitivity
$\mathbf{B}^k$	Quasi-Newton approximation of the Hessian at the $k^{th}$ iteration
$B_{i,p}(u)$	Bézier basis function
$c$	Specific heat, $J/kg\ K$
$c_i(\Phi)$	$i^{th}$ constraint
$C$	Objective function scaling parameter
$\mathbf{C}(u)$	Parametric vector defining the enclosure geometry
$C_{cond}$	Non-dimensional conduction coefficient, Eq. (5.29)
$C_{conv}$	Non-dimensional convection coefficient, Eq. (5.30)
$C_{trans}$	Non-dimensional sensible energy storage coefficient, Eq. (5.31)
$d\mathbf{A}(\Phi)$	Coefficient matrix for first-order temperature sensitivities
$dF_{i-stripj}$	View factor from $u_i$ to an infinitely long strip element centered at $u_j$
$\mathbf{e}_i$	Unit vector in the $i^{th}$ direction

$\mathbf{e}(u, u', \Phi)$	Unit vector at $u$ , pointing at $u'$
$E[]$	Expectation function
$E$	Refinement error
$\%EI$	Energy imbalance
$E_{bi}$	$\sigma T_i^4$ [W/m <sup>2</sup> ]
$\mathcal{E}_{ij}(\Phi)$	Exchange factor between the $i^{th}$ and $j^{th}$ elements
$F(\Phi)$	Objective function
$g(u)$	See Eq. (2.3) and (2.4)
$\tilde{\mathbf{g}}(\Phi)$	Finite difference approximation of the gradient vector, Eq. (4.6)
$h^k$	Interval used in finite difference approximation at the $k^{th}$ iteration
$J(u)$	Surface discriminant, Eq. (3.10)
$k(u_i, u_j)$	Kernel of radiosity equation
$L_c$	Characteristic length, $m$
$\mathbf{n}(u, \Phi)$	Unit normal vector at $u$
$N$	Number of discrete surface elements
$N_b$	Number of bundles used in Monte Carlo simulation
$N_{DS}$	Number of surface elements on the design surface
$N_t$	Number of time steps
$N_{i,p}(u)$	Basis function
$\mathbf{p}^k$	Search direction at the $k^{th}$ iteration
$\mathbf{P}_i$	$i^{th}$ control point

$P(u)$	$x$ -component of $\mathbf{C}(u)$
$q_{cond}(u)$	Heat transfer into an infinitesimal wall element due to conduction from neighboring wall elements, $W/m^2$
$q_{conv}(u)$	Heat transfer into an infinitesimal wall element due to convection from a surrounding fluid, $W/m^2$
$q_o(u)$	Radiosity, $W/m^2$
$q_s(u)$	Heat flux, $W/m^2$
$q_{rad}(u)$	Net radiant heat flux entering a wall element, $W/m^2$
$q'''$	Volumetric heat generation, $W/m^3$
$Q(u)$	$y$ -coordinate of $\mathbf{C}(u)$
$Q_{t-1 \rightarrow t}^{added}(\Phi)$	See Eq. (5.2)
$Q_{t-1 \rightarrow t}^{target}(\Phi)$	See Eq. (5.3)
$Q_{0 \rightarrow tf}^{deficit}(\Phi)$	See Eq. (5.4)
$\mathbf{r}$	Position vector
$s(u_i, u_j, \Phi)$	$C(u_j, \Phi) - C(u_i, \Phi)$
$s^k$	$\Phi^{k+1} - \Phi^k$
$t$	Time, $s$
$T(u)$	Temperature, $K$
$u$	Parameter for enclosure representation
$w_i$	$i^{th}$ singular value

$\mathbf{x}$	Solution vector
$\mathbf{x}'$	First-order sensitivity solution vector
$\mathbf{x}''$	Second-order sensitivity solution vector
$\mathbf{y}^k$	$\nabla F(\Phi^{k+1}) - \nabla F(\Phi^k)$

### Greek Characters

$\alpha$	Step size
$\beta_{ij}$	Blockage factor between the $i^{th}$ and $j^{th}$ elements
$\delta$	Wall thickness, $m$
$\delta$	Residual vector
$\delta_1(\Phi)$	Sampling error in $\tilde{F}(\Phi)$ , Eq. (4.2)
$\delta_{2,p,k}(\Phi^k, h^k)$	Bias error in $\tilde{g}(\Phi)$ , Eq. (4.7)
$\Delta A_i$	Area of the $i^{th}$ element
$\varepsilon$	Emissivity
$\Phi$	Set of design parameters
$\gamma$	Scaling parameter for multi-objective functions
$\Gamma$	Variance-covariance matrix
$\kappa$	Thermal conductivity, $W/m\ K$
$\rho$	Density, $kg/m^3$
$\rho_s$	Specular reflectivity

$\rho_d$	Diffuse reflectivity
$\sigma$	Stefan-Boltzmann constant, $5.67 \times 10^{-8} W/m^2 K^4$
$\sigma_{i,ave}(\Phi)$	Sample standard deviation
$\sigma_i^2(\Phi)$	Sample variance
$\sigma_{ij}^2(\Phi)$	Sample covariance
$\tau$	Non-dimensional time
$\xi$	Vector of random parameters

### Superscripts and Subscripts

$DS$	Design surface
$HS$	Heater surface
$i, j$	Surface element indices
$k$	Iteration number
$p, q$	Design parameter indices
$RS$	Refractory surface
$s$	Scaling parameter
$target$	Desired distribution over the design surface
$*$	Optimal set of design parameters
$0$	Initial set of design parameters
$\wedge$	Non-dimensional parameter
$\sim$	Parameter contains statistical uncertainty

# **Chapter 1:**

## **Introduction**



## 1.1 Motivation

The design of radiant enclosures is a commonly encountered problem in the area of thermal engineering. In many cases, the radiant enclosure is part of a fabrication process where the radiant enclosure acts to thermally treat a product. Such enclosures are found in many different industrial settings; examples include baking ovens in the food industry, rapid thermal processing chambers used in microchip fabrication, annealing furnaces used to heat treat steel, and paint-drying chambers on automotive assembly lines.

The radiant enclosures used in these applications usually conform to a generic type, as shown in Fig. 1.1. The heaters, located on a *heater surface*, are used to heat the products that are placed on a *design surface*. The objective of the design problem, then, is to find the enclosure geometry and heat flux distribution over the heater surface that produces the desired temperature and heat flux distribution over the design surface. (In most cases, a uniform temperature and heat flux distribution is desired over the design surface in order to achieve product homogeneity.) These processes are also often time dependent; for example, a commonly encountered problem is to determine the proper transient heater settings that will heat the product according to a desired temperature history. The main difficulty when designing these enclosures is that, even though both the desired temperature and heat flux distributions over the design surface are known, for the problem to be mathematically well posed only one thermal boundary condition can be specified over each surface. In this sense, the problem (like most design problems) is an *inverse* problem.

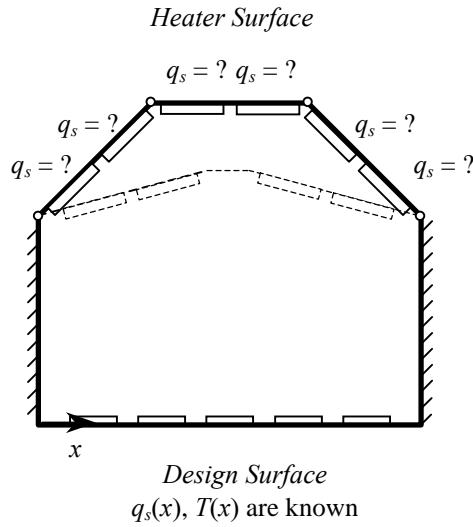


Figure 1.1: Example of a radiant enclosure design problem.

Most inverse problems are fundamentally ill posed; unlike forward problems, which usually have a unique mathematical solution, inverse problems often have many possible solutions; on the other hand, it is also possible that the problem has no solution. In this application, there may be several different enclosure geometries and heat flux distributions that produce both the desired temperature and heat flux distribution over the design surface within an acceptable tolerance, although it is also possible no enclosure configuration exists that would produce the desired design outcome.

This situation is more clearly seen by comparing the systems of linear equations that are often produced by discretizing well-posed and ill-posed mathematical problems, as part of a solution procedure. Well-posed problems produce well-conditioned systems of linear equations, where the number of equations equals the number of unknowns. In these cases there exists one unique solution to the problem, which can be found using

traditional linear algebra techniques. On the other hand, ill-posed problems often produce linear systems where the number of equations does not equal the number of unknowns. If there are more unknowns than equations, then the system is *under-determined* and an infinite set of solutions satisfies the linear system. If there are more equations than unknowns, the system is said to be *over-determined* and there is not likely to exist any solution that satisfies the linear system without a residual vector.

Because inverse problems are very difficult to solve, until recently radiant enclosures were designed exclusively using a forward “trial-and-error” methodology. In this approach, the designer specifies only one of the known design surface boundary conditions, either the temperature or the heat flux, and then guesses an enclosure configuration that might produce the other unspecified boundary condition over the design surface. This transforms the ill-posed inverse problem into a well-posed forward problem having only one explicit solution, which can be found using traditional linear algebra techniques. The unspecified design surface boundary condition is then compared with the desired distribution; if the desired distribution does not match the realized distribution (as is usually the case), the designer modifies the enclosure configuration according to his or her experience and intuition and repeats the analysis. In the example shown in Fig. 1.2, the designer first specifies the temperature distribution over the design surface, and then repeatedly tries different heat flux distributions over the heater surface until one is found that produces a heat flux distribution over the design surface that matches the desired distribution within an acceptable tolerance. The complex nature of radiant heat transfer between the enclosure surfaces makes an intuitive understanding of

the heat transfer physics elusive, however, and consequently it is very difficult for the designer to know how best to adjust the enclosure configuration at each design iteration. Accordingly, many iterations are required to identify a suitable solution, and while the final solution may be satisfactory, it is rarely optimal.

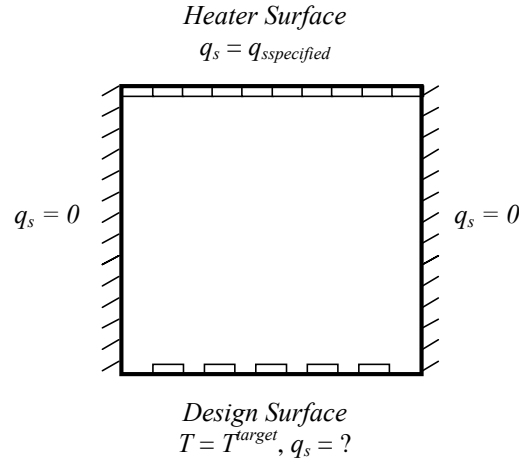


Figure 1.2: Boundary conditions for the forward design methodology.

Two alternative approaches have been developed that overcome the shortcomings of the forward “trial-and-error” design technique. In the inverse design methodology, the inverse problem is posed *explicitly*, where both the temperature and heat flux distributions over the design surface are enforced, while the heat flux distribution over the heater surface remains unspecified; an example of this is shown in Fig. 1.3. The resulting problem is mathematically ill-posed and produces an ill-conditioned set of equations when it is discretized. As previously noted, traditional linear algebra techniques such as Gaussian elimination and LU decomposition either fail completely to solve the set of equations or produce solutions that are noisy and often non-physical. (In

the case of radiant enclosure design, non-physical solutions involve negative emissive power.) Rather, special *regularization* techniques must be employed to solve the linear system; these methods work by finding smooth solutions that satisfy the set of equations with a small error. Smoother solutions are usually easier to implement but also produce a larger residual when they are substituted into the system of equations. Therefore, most regularization algorithms include a heuristic parameter that allows the user to choose between solutions of varying degrees of smoothness and accuracy. These techniques can usually identify a solution to the inverse problem in very few iterations, although it is often in a form that would be difficult to implement in a practical setting. Also, inverse design methods have not as of yet been applied to design radiant enclosure geometry and have largely been restricted to calculating the heat flux distribution from radiant heaters that produce the desired temperature and heat flux distributions over the design surface.

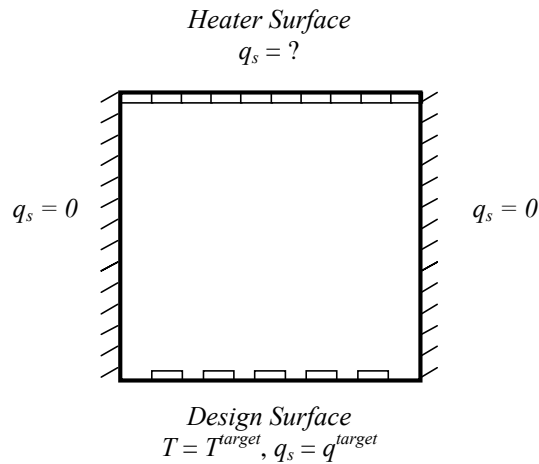


Figure 1.3: Boundary conditions for the inverse design methodology.

This work presents five optimization methodologies that can be used to solve different types of radiant enclosure design problems. Like the forward methodology, optimization design methodologies solve the inverse design problem by repeatedly analyzing trial designs and then making adjustments based on the results of the analysis. Instead of relying solely on the designer's intuition and experience to make the modifications, however, optimization methodologies work by systematically minimizing an *objective function*,  $F(\Phi)$ , defined in such a way that the minimum of the objective function corresponds to the desired design outcome. The objective function is dependent on a set of *design parameters*,  $\Phi$ , which specify the design configuration; in this application, the design parameters control the enclosure geometry and the heat flux distribution over the heater surface. The goal, then, is to identify the set of design parameters  $\Phi^*$  that minimizes the objective function, i.e.  $F(\Phi^*) = \text{Min}[F(\Phi)]$ , which in turn corresponds to the optimal design configuration.

Whereas both thermal boundary conditions are explicitly enforced throughout the solution procedure in the inverse methodology, in the optimization methodology only one of the desired thermal boundary conditions on the design surface is enforced while the remaining one is used to define the objective function. (The inverse problem is solved in its well-posed *implicit* form, since only one of the two desired thermal boundary conditions is specified over the design surface.) Most often, the objective function is the variance of the unspecified boundary condition distribution corresponding to a given set of design parameters from the desired distribution, evaluated at discrete locations over the design surface. In the example shown in Fig. 1.4, the temperature distribution is

enforced over the design surface and the design parameters control the heat flux distribution over the heater surface. The resulting forward problem is solved for the heat flux distribution over the design surface, which in turn is used to evaluate  $F(\Phi)$ .

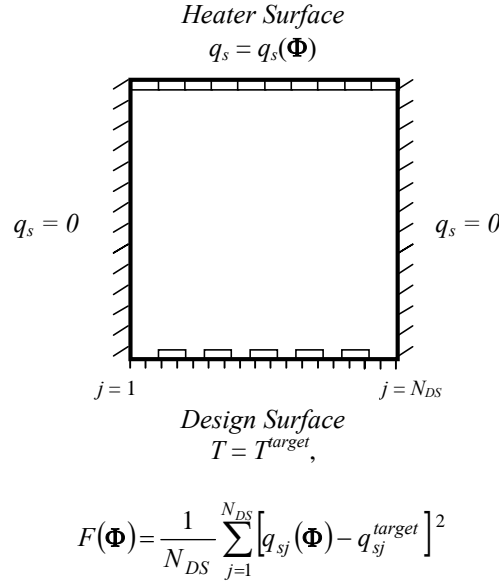


Figure 1.4: Boundary conditions for the optimization design methodology.

Since the objective functions encountered in radiant enclosure design problems are usually nonlinear functions of  $\Phi$ , minimization is carried out using *nonlinear programming* (NLP) algorithms that work by making intelligent changes to the design parameters at each iteration based on the local topography of  $F(\Phi)$  until  $F(\Phi^*)$  is reached. Consequently, far fewer iterations (and consequently far less design time) is required compared to the forward design methodology, and the quality of the final solution found using the optimization design methodology is almost always better than one found through the forward design methodology.

The optimization methodology also has several important advantages over the inverse design methodology. First, while the inverse methodology requires the design problem to be written in an explicit form so it can be solved using a regularization technique, the optimization design methodology solves the problem in its “natural” implicit form. Consequently, the optimization design methodology usually requires less time and mathematical expertise on the part of the designer to set up and solve the problem compared to the inverse design methodology. This aspect also renders the optimization methodology applicable to a wider range of design problems; in particular, enclosure geometry design problems are very difficult to pose in an explicit form, but are more easily solved in their implicit form using optimization techniques.

Secondly, unlike the inverse design methodology, it is relatively easy to impose design constraints in the optimization methodology, by forcing the design parameters to lie within a fixed domain called the “feasible region.” Design constraints help ensure that the solution found through the optimization methodology can be easily implemented in a physical setting. In the context of radiant enclosure design problems, constraints can be used to restrict the heat flux distribution over the heater surface to be greater than zero and less than the maximum heater output. In the case of enclosure geometry design, constraints also ensure that the size of the enclosure is less than some user-defined maximum dimensions, and that the enclosure remains unobstructed (meaning that the interior of the enclosure surface is strictly concave), which greatly simplifies the analysis. (Björck, 1992, lists several ways to accommodate simple constraints during the



regularization procedure, but these techniques are quite complex and have yet to be applied to solve the inverse radiant enclosure design problem.)

## **1.2 Literature Review**

In order to understand the state of radiant enclosure design, it is instructive to examine the development of techniques for solving inverse radiation heat transfer problems presented in the literature. Both inverse and optimization analysis techniques were initially developed as tools used in experimental settings to infer properties that could not be measured directly, based on other measured experimental data, and were later adapted to solve inverse design problems. A more detailed discussion of the development of techniques for solving inverse heat transfer problems is provided by Özisik and Orlande (2000), Beck et al. (1985), and Alifanov et al. (1995).

### **1.2.1 Use of Inverse and Optimization Techniques in Experimental Settings**

Inverse heat transfer problems were first studied in the space programs of the 1950s, with a goal of developing algorithms that could estimate the temperature and ablation rates of spacecraft surfaces exposed to atmospheric heating during reentry based on temperature measurements taken on the undersides of these surfaces. Tikhonov (1975) proposed a regularization method for solving the system of ill-conditioned equations that arises from the explicit inverse conduction problem, which bears his name. Other regularization approaches were later developed, including those by Beck et al. (1985).

Optimization techniques were also first applied in the area of experimental heat transfer. They were, in fact, first used by Alifanov (1974) as regularization tools for solving inverse conduction problems in their explicit form. In this technique, the ill-conditioned system of equations is used to form an objective function with a global

minimum corresponding to a solution that satisfies the equations with no residual error. Each successive iteration of a gradient-based minimization scheme increases the solution accuracy at the expense of smoothness; accordingly, the minimization procedure is terminated once a solution that has a small enough residual and a sufficient degree of smoothness is identified, which may occur before a local minimum is reached. (Iterative regularization through optimization is described more thoroughly in Appendix B.)

Only later were optimization methods used to solve inverse heat transfer problems implicitly. In most cases, they were again used to infer the distribution of a quantity that cannot be measured directly; this is done by first defining an objective function as the difference between the distribution of an experimentally measured property and the distribution of the same property estimated through a numerical model where the only variable is the distribution of the unmeasurable quantity. This distribution is then inferred by minimizing this objective function. For example, Huang and Chen (1999) estimated the thickness of an ice layer deposited on a cold pipe immersed in a moist stream of air by minimizing an objective function defined as the difference between the pipe surface temperature distribution calculated using a numerical model where the ice layer thickness is the only variable and the temperature distribution determined experimentally using thermocouples. Similar techniques were employed by Liu and Özisik (1996) to estimate the inlet temperature profile of fluid flowing between parallel plates based on temperature measurements made inside the channel, and by Liu (2000), who estimated the temperature profile and absorption coefficient within an absorbing-emitting media based on exit radiation intensities at the boundary surfaces.

### 1.2.2 Use of Inverse Techniques to Solve Design Problems

While regularization techniques have been widely used in experimental settings, only recently have they been applied to solve inverse design problems involving radiant enclosures. The first type of problem studied was the inverse boundary condition (or load estimation) problem, where the goal is to identify a heat flux distribution over the heater surface that produces the desired heat flux and temperature distributions over the design surface, as shown in Fig. 1.5.a. Oguma and Howell (1995) were the first to study such a problem, and applied an inverse Monte Carlo technique to treat enclosures consisting of diffuse, black walls containing a transparent medium. Inverse design methodologies were next developed to solve inverse design problems involving enclosures containing diffuse-gray surfaces; Harutunian et al. (1995) solved the ill-conditioned set of linear equations using the modified truncated singular value decomposition method (MTSVD), which was later extended by Morales et al. (1996) to treat radiant enclosures containing isothermal participating media. Jones (1999) modified the method presented by Harutunian et al. (1995) by adding an extra equation based on energy conservation to the ill-conditioned matrix equation, which was then solved using Tikhonov regularization. França and Goldstein (1996) developed an inverse design methodology to solve design problems involving non-isothermal participating media, where the governing equations were first formed using Hottel's zonal method, and then solved by performing incomplete Jacobi and Gauss-Seidel iteration. França et al. (2001) later extended this technique to solve a radiant enclosure design problem involving conduction, convection, and radiation heat transfer within a participating medium. The

nonlinear set of equations was first linearized in such a way that the coefficient matrix is constant while the elements of the right-hand-side vector are nonlinear functions of the solution vector. The problem was then solved using truncated singular value decomposition (TSVD) regularization, by first decomposing the coefficient method into its singular values, and then finding the solution vector by performing repeated back-substitution iterations.

A second type of radiant enclosure design problem is shown in Fig. 1.5.b, where the objective is to determine the distribution of a radiant energy source inside a participating medium that results in the desired heat flux and temperature distributions over the design surface. Kudo et al. (1995) first developed an inverse design technique for estimating the optimal location of sources having specified power inputs, which used Monte Carlo to carry out the radiation heat transfer analysis and singular value decomposition to solve the ill-conditioned set of equations. França et al. (1999) solved the complement of that problem, with the objective of determining the heat generation rate from radiant sources at fixed locations within the participating medium that produce the required conditions over the design surface.

As previously mentioned, many radiant enclosure design problems are transient in nature, with the objective of determining the time-dependent heater settings that heat the design surface according to a specified temperature history, such as the one shown in Fig. 1.5.c. Ertürk et al. (2002) calculated the heater settings in a roll-through furnace having surfaces with directionally-dependent optical properties and multimode heat transfer effects. The problem was solved by first discretizing the time domain, linearizing the

resulting nonlinear set of equations, and then solving the inverse problem by applying a fixed number of conjugate gradient regularizing iterations at each time step. Gwak and Masada (2003) later solved the same problem using a nonlinear control algorithm coupled with a regularization filter.

### 1.2.3 Use of Optimization Techniques to Solve Design Problems

Optimization techniques have only recently been applied to solve inverse design problems involving radiant enclosures. Daun et al. (2003a) were the first to use gradient-based optimization techniques to solve an inverse boundary condition problem similar to the one shown in Fig. 1.5.a., for a 2-D enclosure with diffuse-gray walls containing a non-participating medium. The heat flux distribution and sensitivities were calculated using the infinitesimal-area method, and the optimization was carried out using Newton's method. Hosseini Savari et al. (2003a) presented a method of designing 2-D enclosures containing a conducting and radiating medium where the objective function and sensitivities were calculated with the discrete transfer method and the objective function was minimized using the conjugate gradient method. This method was later extended to treat 3-D enclosures containing participating media (Hosseini Savari et al., 2003b.)

Transient radiant enclosure design problems like the one shown in Fig. 1.5.b have also been solved using the optimization design methodology. Norman (1992) and Cho and Gyugyi (1997) used optimization to determine an *a priori* guess of the transient heater settings used in RTP furnaces, enabling the use of high gain control algorithms under operating conditions. Fedorov et al. (1998) solved for heater settings in a roll-

through materials processing furnace, using the Levensberg-Marquardt optimization method to carry out the minimization. Daun et al. (2003b) presented an optimization methodology based on infinitesimal-area analysis for solving transient radiant enclosure design problems, where the objective function is minimized using the quasi-Newton method.

Relatively little work thus far has been carried out on the inverse radiant enclosure geometry design problem. Ashdown (1994) described a general optimization methodology that uses genetic algorithms to optimize the geometry of light boxes used in illumination applications. Daun et al. (2003c) presented a methodology for optimizing the geometry of diffuse walled enclosures based on parametric representation, where the objective function and sensitivities are calculated using the infinitesimal-area analysis and the minimization is carried out using either the steepest-descent, quasi-Newton, or Newton's method. Daun et al. (2003d) later presented another technique for designing enclosures containing specularly-reflecting surfaces, where the radiation heat transfer is solved using Monte Carlo and the Kiefer-Wolfowitz method is used to carry out the stochastic optimization. There has also been some recent work on geometric optimization of radiating fin arrays motivated by extraterrestrial cooling applications (Krishnaprakas and Badari Narayana, 2003, for example), although for the most part the optimization is either univariate, where the problem has only one design parameter, or is carried out using parametric analysis instead of more sophisticated nonlinear programming techniques.

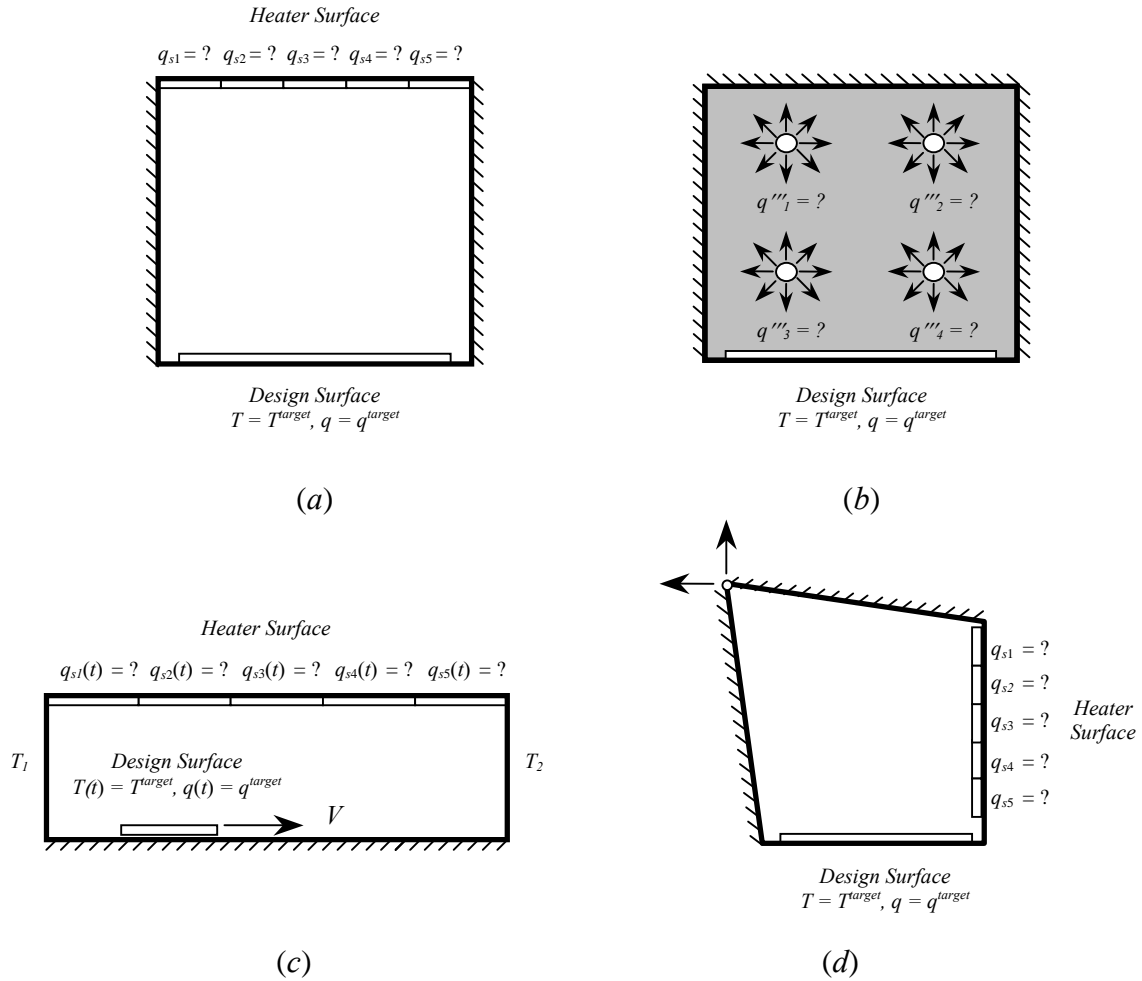


Figure 1.5 Types of radiant enclosure design problems solved using inverse and optimization design methodologies in the literature: (a) inverse boundary condition estimation, (b) inverse source estimation, (c) transient problems, and (d) geometry design problems.



### 1.3 Dissertation Outline

This dissertation is divided into six chapters; chapters two through five describe optimization methodologies for different types of radiant enclosure problems, and conclusions and suggestions for future work are presented in chapter six.

Chapter two serves as an introduction to the forward, inverse, and optimization design methodologies used to solve radiant enclosure design problems. In this chapter the design problem is in its simplest form, where the objective is to find the heater settings that produce the desired heat flux and temperature distribution over the design surface in a radiant enclosure having a fixed geometry, composed of diffuse-gray walls containing a transparent medium. The inverse and optimization design methodologies are compared by using them to solve an example problem; the inverse methodology solves the ill-conditioned set of linear equations using TSVD regularization, while the optimization methodology uses Newton's method to minimize the objective function coupled with an active set method to enforce design constraints.

Chapter three extends the optimization methodology presented in chapter two to design the geometry of diffuse-walled radiant enclosures. This is done by first representing the surface geometry with Bézier curves and then using control points of the curves as the design parameters in the optimization. Newton's method coupled with an active set method is again used to carry out the objective function minimization.

A methodology for designing the geometry of radiant enclosures containing specular surfaces is presented in chapter four. The radiation heat transfer is modeled with

the Monte Carlo analysis method, and the Kiefer-Wolfowitz method is used to carry out the objective function minimization.

Finally, chapter five presents two optimization design methodologies for solving transient and multimode radiant enclosure design problems, where the objective is to identify transient heater settings that heat the design surface according to a desired temperature history. The first optimization methodology considers the effects of conduction through the enclosure walls, convection between the walls and surrounding fluids, and sensible energy storage, and can therefore be considered “generic” in the sense that a wide range of problems can be treated using this methodology. A second methodology is then presented for the special case where there is no conduction or convection heat transfer, and the design surface is the only part of the enclosure that stores sensible energy. It can be shown that in this case the heat transfer model reduces to a set of linear equations, which greatly simplifies the analysis. In both of these methodologies, the optimization is carried out using a quasi-Newton minimized coupled with an active set method to enforce design constraints.

The purpose of this work is to present optimization methodologies that can be used to solve different fundamental types of radiant enclosure design problems in as clear a manner as possible. Accordingly, the enclosure walls are assumed to have wavelength-independent optimal properties in each case in order to clarify the analyses, although this restriction could be relaxed with little difficulty if so required.

**Chapter 2:**  
**Heater Setting Optimization for Radiant**  
**Enclosures Containing Diffuse-Gray Surfaces**

## 2.1 Introduction

This chapter presents forward, inverse, and optimization design methodologies that are used to solve radiant enclosure design problems where the enclosure geometry is fixed and the heater settings that control the heat flux distribution over the heater surface are unknown. The objective of these design problems is to find the heater settings that produce the desired heat flux and temperature distributions over the design surface. This is the simplest type of radiant enclosure design problem, and is found in many different industrial settings. Often, these problems involve enclosures used in continuous fabrication processes in which the enclosure geometry and temperature and heat flux distributions can be modeled as constant with respect to time; such examples include industrial ovens that bake food products on a conveyor belt and annealing furnaces used in continuous casting processes.

These design problems have traditionally been solved using a forward “trial-and-error” methodology. In this approach, only one of the two desired distributions (either temperature or heat flux) over the design surface is specified. The designer first guesses a set of heater settings, and then evaluates the quality of this guess by solving the resulting well-posed problem for the distribution of the unspecified boundary condition over the design surface. If the resulting distribution does not match the desired distribution, which is usually the case, the designer adjusts these heater settings using only his or her intuition and experience, and then repeats the analysis. This process continues until the heat flux and temperature distributions over the design surface match the desired distributions within an acceptable tolerance. (Since only one of the desired boundary conditions is enforced and the other is found through iteration, the inverse

problem is said to be in its *implicit form* here.) This technique usually requires many iterations and a lot of design time; furthermore, this solution is usually far from optimal

The shortcomings of this technique prompted the application of inverse design methodologies to solve this problem. In this approach, both the temperature and the heat flux distributions over the design surface are enforced explicitly throughout the solution procedure. (Hence, the inverse problem is solved in its *explicit form*.) When both boundary conditions are enforced on the design surface, the governing equations are ill-posed and the corresponding set of ill-conditioned linear equations must be solved using regularization methods. Inverse methodologies require very few iterations to find heater settings that produce temperature and heat flux distributions over the design surface that are nearly identical to the desired distributions. Oguma and Howell (1995) were the first to apply an inverse design methodology to solve a radiant enclosure design problem, which used an inverse Monte Carlo analysis to find the heater settings in an enclosure with black and diffuse surfaces. Harutunian et al. (1995) developed an inverse design methodology for enclosures composed of diffuse-gray walls containing a non-participating medium, where the ill-conditioned set of equations is formed using discrete configuration factors and solved using modified truncated singular value decomposition. This technique was later extended to treat radiant enclosures containing participating media (Morales et al., 1996, França and Goldstein, 1996) and problems involving multimode heat transfer (França et al., 2001.) The main drawback to the inverse design methodology, however, is that solutions found through regularization can sometimes be non-physical (having negative emissive powers) or are otherwise difficult to implement. In particular, solutions are often found that involve regions of negative heat flux over the

heater surface; since it is impractical to produce this condition in a physical setting these regions are most often insulated, which degrades the quality of the heat flux and temperature distributions over the design surface.

The optimization design methodology overcomes many drawbacks of the inverse design methodology. As is done in the forward design methodology, the inverse problem is written in its implicit form; only one of the desired design surface boundary conditions (the temperature, by convention) is enforced while the unspecified heat flux distribution is used to define an objective function,  $F(\Phi)$ , in such a way that the minimum of  $F(\Phi)$  corresponds with the desired heat flux distribution over the design surface. The heat flux distribution, and hence  $F(\Phi)$ , are dependent on the heater settings, which are set equal to design parameters contained in  $\Phi$ . Specialized numerical methods are then used to minimize the objective function iteratively by making changes to  $\Phi$  at each iteration based on the local topography of  $F(\Phi)$ . The set of design parameters  $\Phi^*$  that minimizes the objective function correspond to the optimal heater settings.

Although the optimization methodology usually requires more iterations than the inverse methodology to find a solution, the final solution is almost always physical since the inverse problem is solved in its well-posed implicit form. Furthermore, design constraints can be imposed by forcing the design parameters to lie within a specified domain called the feasible region, which are used to ensure that the solution can be implemented in a physical setting. In this example, design constraints account for the operating range of the heaters by forcing the heater settings to lie between zero and the maximum heater output.

This chapter serves as an introduction to the inverse and optimization design methodologies, and is based on a recent work by Daun et al. (2003a). Both techniques presented here are based on the infinitesimal area method developed by Daun and Hollands (2001), and are in a form appropriate for the analysis of two-dimensional radiant enclosures. The inverse design problem is solved using truncated singular value decomposition (TSVD), while the optimization is carried out using Newton's method coupled with an active set method that enforces the design constraints. The forward methodology is first presented, followed by the inverse and optimization methodologies, respectively. The inverse and optimization methodologies are then demonstrated and compared by using them to solve for the unknown heater settings in a 2-D radiant enclosure design problem.

## 2.2 Forward Design Methodology

The forward design methodology is based on the infinitesimal-area technique presented by Daun and Hollands (2001). In this approach, the designer first specifies the enclosure geometry and boundary conditions, including a guess of the heat flux distribution over the heater surface and one of the designed boundary conditions over the design surface. (Here, it is assumed to be the temperature distribution.) Once this is done, the integral equations governing the radiosity distribution are formed based on a parametric representation of the enclosure, and then solved numerically. Finally, the unknown boundary condition over the design surface is found by post-processing the radiosity solution. The designer compares this result to the desired distribution and modifies the heater settings accordingly. This process is repeated until a satisfactory solution is identified.

The first step in this analysis is to find a suitable parametric representation of the radiant enclosure; for 2-D enclosures, it is sufficient to identify a parametric representation of the enclosure cross section,

$$\mathbf{r} = \mathbf{C}(u) = \{P(u), Q(u)\}^T, \quad a \leq u \leq b, \quad (2.1)$$

where  $\mathbf{r}$  is the position vector with its tail at the origin and tip on the enclosure surface and  $\mathbf{C}(u) = \{P(u), Q(u)\}^T$  is a vector function of parameter  $u$  with components  $P(u)$  and  $Q(u)$  in the  $x$  and  $y$  directions, respectively. By allowing  $u$  to range over its entire domain,  $a \leq u \leq b$ , the position vector carves out the enclosure cross section in the  $x$ - $y$  plane, as shown in Fig. 1. (By convention, the enclosure is usually parameterized in a counterclockwise manner so that the normal vector points into the enclosure.) The boundary conditions are also expressed parametrically; the emissivity distribution,  $\epsilon(u)$  is



known, and either the heat flux,  $q_s(u)$ , or the temperature,  $T(u)$ , is known at every location over the enclosure surface. In particular, it is assumed that  $T^{target}(u)$  is specified over the design surface and an initial guess of the heater settings,  $q_s(u)$ , is imposed over the heater surface.

Once the enclosure is represented parametrically, the equation governing the radiosity distribution is formed. The radiosity distribution is governed by a Fredholm integral equation of the second-kind,

$$q_o(u) = b(u) + g(u) \int_a^b q_o(u') k(u, u') du', \quad (2.2)$$

where if the temperature is specified at  $u$ ,

$$b(u) = \varepsilon(u) E_b(u), \quad g(u) = 1 - \varepsilon(u), \quad (2.3)$$

with  $E_b(u) = \sigma T^4(u)$ , or if the heat flux is specified at  $u$ ,

$$b(u) = q_s(u), \quad g(u) = 1. \quad (2.4)$$

The kernel of Eq. (2.1),  $k(u, u')$ , equals the differential configuration factor from a point on the enclosure surface at  $u$  to an infinitely long strip of infinitesimal width centered at  $u'$ , divided by  $du'$ , and is derived from the parametric functions that represent the enclosure cross section. (The kernel is defined algebraically in Chapter 3.)

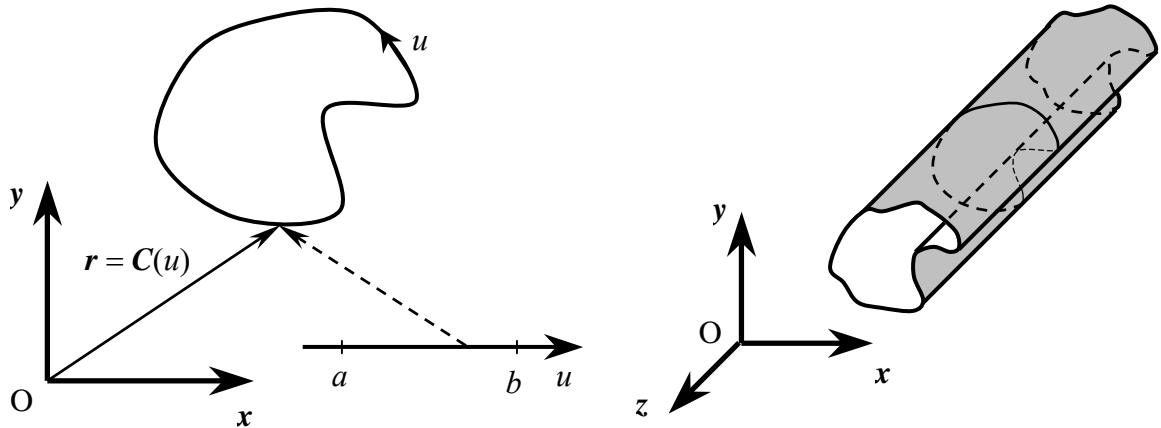


Figure 2.1: Parametric representation of the enclosure cross-section.

Analytical solutions to Fredholm integral equations of the second kind are usually not tractable, so Eq. (2.2) must be solved numerically. The first step is to discretize the parametric domain into  $N$  elements, the  $i^{th}$  element centered at  $u_i$  and having a width  $\Delta u_i$ . Each of these discrete elements corresponds to an infinitely long strip element having a finite width on the enclosure surface, as shown on Fig. 2.2. The integral in Eq. (2.2) is then rewritten as the sum of  $N$  integrals, with the domain of the  $j^{th}$  integral defined by  $u_j - \Delta u_j/2 < u < u_j + \Delta u_j/2$ ,

$$q_o(u) = b(u) + g(u) \sum_{j=1}^N \int_{u_j - \Delta u_j/2}^{u_j + \Delta u_j/2} q_o(u') k(u, u') du' \quad (2.5)$$

By assuming a uniform radiosity distribution over each element,  $q_o(u')$  can be extracted from the integral. Carrying out the remaining integrations results in

$$q_{oi} = b_i + g_i \sum_{j=1}^N q_{oj} \beta_{ij} dF_{i-stripj}, \quad (2.6)$$

where  $q_{oi} = q_o(u_i)$ ,  $b_i = b(u_i)$ ,  $g_i = g(u_i)$ , and  $dF_{i-stripj}$  is the differential configuration factor  $u_i$  to the strip element centered on  $u_j$ . (The calculation of  $dF_{i-stripj}$  is discussed in the next chapter.) The blockage factor,  $\beta_{ij}$  is a binary term that equals zero if the path between  $u_i$  and  $u_j$  is blocked by an obstructing surface, and is otherwise equal to unity. It is found either through analytical geometry or by numerically using a ray-tracing technique.

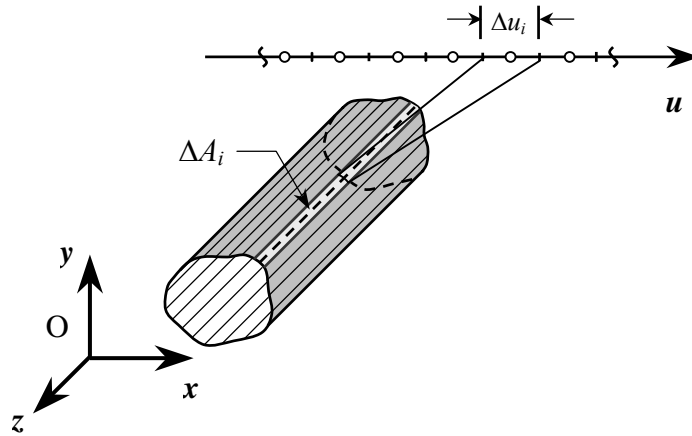


Figure 2.2: Discretization of the parametric domain.

Writing Eq. (2.6) for all the elements results in a system of  $N$  equations containing  $N$  unknowns, which can be rewritten as a matrix equation,

$$\mathbf{A} \mathbf{x} = \mathbf{b}, \quad (2.7)$$

where  $\mathbf{x} = \{q_{o1}, q_{o2}, \dots, q_{oN}\}^T$ ,  $\mathbf{b} = \{b_1, b_2, \dots, b_N\}^T$ , and matrix  $\mathbf{A}$  is defined by

$$A_{ij} = \begin{cases} 1 & i = j \\ -g_i \beta_{ij} dF_{i-stripj} & i \neq j \end{cases}. \quad (2.8)$$

Matrix  $\mathbf{A}$  is usually well-conditioned, and Eq. (2.7) can be solved using any standard linear algebra technique. Once the radiosity distribution is known, the unspecified boundary conditions are found by post-processing. In particular, the heat flux distribution over the design surface is found by

$$q_{si} = q_{oi} - \sum_{j=1}^N q_{oj} \beta_{ij} dF_{i-stripj}, \quad (2.9)$$

where  $q_{si} = q_s(u_i)$ . If the resulting heat flux distribution over the design surface does not match the desired heat flux distribution, the designer adjusts the heater settings according to his or her experience and intuition and then repeats the analysis. This procedure continues until the heat flux distribution over the design surface matches the desired distribution,  $q_s^{target}(u)$  within an acceptable tolerance, which typically requires many manual iterations to accomplish.

### 2.3 Inverse Design Methodology

In contrast to the forward design methodology where only one thermal boundary condition is specified over each surface, the inverse design methodology enforces both the temperature and heat flux distributions over the design surface, while no boundary condition is specified over the heater surfaces.

Suppose that the enclosure surface is parameterized so that  $a \leq u < b$  corresponds to the heater surface,  $b \leq u < c$  corresponds to the surfaces where either  $q_s(u)$  or  $E_b(u)$  is known, and  $c \leq u < d$  corresponds to the design surface. Such a radiant enclosure is shown in Fig. 2.3. (Although there are certain advantages in having  $u$  represent the enclosure surfaces continuously in a counterclockwise order, this is not required.) Since both temperature and heat flux are specified over the design surface, the radiosity distribution is found immediately by

$$q_{oi}^{target}(u) = E_{bi}^{target}(u) - \frac{1 - \varepsilon(u)}{\varepsilon(u)} q_s^{target}(u), \quad c \leq u < d. \quad (2.10)$$

The radiosity distribution over the remaining enclosure surfaces is again governed by Eq. (2.2). In the inverse methodology, however, the integral is rewritten as the sum of two integrals: one with a domain of  $a \leq u < c$ , over which the radiosity is unknown, and the other with a domain  $c \leq u < d$ , where the radiosity distribution over the design surface is known from Eq. (2.10). Over the heater surface, then, the radiosity distribution is governed by

$$q_o(u) - q_s(u) - \int_a^c q_o(u') k(u, u') du' = \int_c^d q_o^{target}(u') k(u, u') du', \quad a \leq u < b, \quad (2.11)$$

while for the remaining enclosure surfaces,

$$q_o(u) - g(u) \int_a^c q_o(u') k(u, u') du' = b(u) + g(u) \int_c^d q_o^{target}(u') k(u, u') du', \quad b \leq u < c. \quad (2.12)$$

where  $b(u)$  and  $g(u)$  are again defined by Eqs. (2.3) or (2.4), depending on if  $u$  lies on a surface where  $T(u)$  or  $q_s(u)$  is specified, respectively. The terms on the left-hand sides of Eqs. (2.3) and (2.4) are unknown while those on the right-hand side are known, since the integrals on the right-hand side can be carried out.

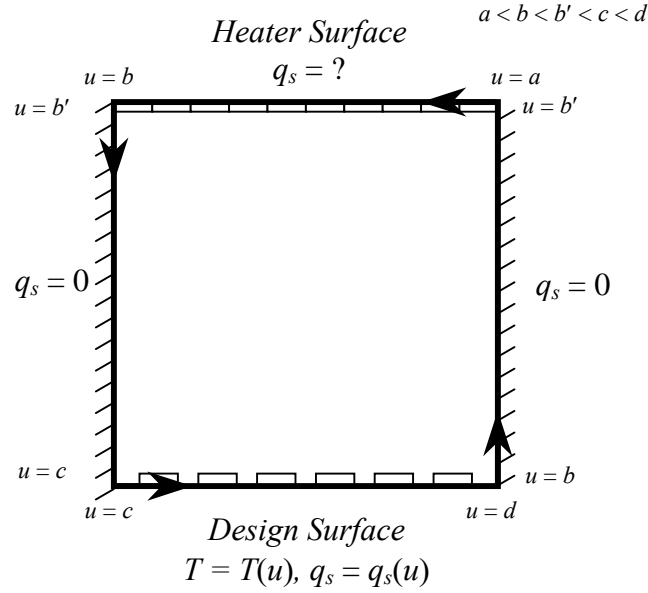


Figure 2.3: Example of an inverse radiant enclosure design problem posed in its explicit form.

At this point, it is useful to compare the governing radiosity equations in the forward and inverse design methodologies. In the forward methodology, the radiosity distribution was governed by a single Fredholm integral equation of the second-kind, Eq. (2.2), which most often has a unique solution. In contrast, Eqs. (2.10) and (2.11) are not “true” Fredholm integral equations of the second-kind, since both the radiosity and heat flux distributions are unknown over parts of the parametric domain. Because of this, the system formed by these two equations is under-specified, and an infinite set of solutions

could satisfy these equations. This is more clearly seen by writing the governing radiosity equation over the design surface,

$$q_o^{target}(u) - b(u) - g(u) \int_c^d q_o^{target}(u') k(u, u') du' = g(u) \int_a^c q_o(u') k(u, u') du', \quad c \leq u < d, \quad (2.13)$$

where  $b(u)$  and  $g(u)$  are defined using either Eq. (2.3) or (2.4), and every other term is known except for the radiosity distribution  $q_o(u')$  for  $a \leq u' < d$ . Equation (2.13) is a Fredholm integral equation of the first-kind, which differs from Fredholm equations of the second-kind in the sense that there are no unknowns outside of the integrals. In most cases, an infinite set of radiosity distributions could be proposed that satisfy Eq. (2.13); it is this feature of the governing equations that renders the inverse problem ill-posed in its explicit form. (Note that since the radiosity distribution over the design surface is known, Eq. (2.13) is superfluous and is included only for demonstration purposes.)

Analytical solutions to Eqs. (2.11) and (2.12) are usually not tractable, so numerical techniques must again be employed to solve for the radiosity distribution. As before, the parametric region is discretized to form  $N$  surface elements. Let elements  $i = 1 \dots k$  lie on the heater surface, elements  $i = k + 1 \dots m$  lie on surfaces where either  $q_s(u_i)$  or  $T(u_i)$  is known, and elements  $i = m + 1 \dots N$  lie on the design surface. By assuming a uniform radiosity over each surface element and performing the same steps as before, the governing equations can be rewritten as

$$q_{oi} - q_{si} - \sum_{\substack{j=1 \\ j \neq i}}^m q_{oj} \beta_{ij} dF_{i-stripj} = \sum_{j=m+1}^N q_{oj} \beta_{ij} dF_{i-stripj}, \quad i = 1 \dots k, \quad (2.14)$$

for elements over the heater surface, and

$$q_{oi} - g_i \sum_{\substack{j=1 \\ j \neq i}}^m q_{oj} \beta_{ij} dF_{i-stripj} = b_i + g_i \sum_{j=m+1}^N q_{oj} \beta_{ij} dF_{i-stripj}, \quad i = k + 1 \dots m \quad (2.15)$$

for all other elements.

Writing Eqs. (2.14) and (2.15) for all the elements with unknown radiosities results in a system of  $m$  equations containing  $m + k$  unknowns, the extra  $k$  unknowns due to the unknown heat flux distribution over the heater surfaces. As before, these equations can be expressed as a matrix equation,

$$\mathbf{A} \mathbf{x} = \mathbf{b}, \quad (2.16)$$

where  $\mathbf{A}$  is a rectangular matrix having  $m$  rows and  $m + k$  columns and  $\mathbf{x}$  is a vector containing  $m$  unknown radiosity terms and the  $k$  unknown heat flux terms. Vector  $\mathbf{b}$  is formed from the known radiosity terms on the design surface and the specified boundary conditions (either  $q_{si}$  or  $E_{bi}$ ) from the other surfaces.

Since there are more unknowns than equations, matrix  $\mathbf{A}$  is rank-deficient and admits an infinite number of solutions. Consequently, traditional linear algebra techniques that are normally used to solve matrix equations, such as LU decomposition and Gauss-Seidel, fail when they are applied to solve Eq. (2.16). Instead, regularization techniques are used to select solutions having desirable properties from the infinite set, in general by reducing the norm of  $\mathbf{x}$  at the expense of solution accuracy. These methods include Tikhonov regularization, iterative conjugate gradient regularization, and the truncated singular value decomposition (TSVD) method. Although any of these methods could be used to solve Eq. (2.16), TSVD offers the most insight into the nature of ill-posed problems, and is demonstrated here. (Several other techniques are described and demonstrated in Appendix B.)

The TSVD method is based on the singular decomposition of the  $\mathbf{A}$  matrix,

$$\mathbf{A} = \mathbf{U} \mathbf{W} \mathbf{V}^T, \quad (2.17)$$

where  $\mathbf{U}$  is an orthogonal matrix with  $m$  rows and  $m + k$  columns,  $\mathbf{W}$  is a diagonal matrix with  $m + k$  rows and columns, and  $\mathbf{V}$  is the transpose of an  $(m + k) \times (m + k)$  orthogonal matrix. The diagonal elements of  $\mathbf{W}$ ,  $w_i$ , are the singular values, which are strictly non-negative and are usually presented in monotonically decreasing order. The degree of difficulty associated with solving a linear system of equations can be anticipated by examining the singular values and specifically the *condition number*, which is the ratio of the largest and smallest singular value. If all the singular values are approximately the same magnitude and the condition number is of  $O[1]$ , then  $\mathbf{A}$  is said to be *well-conditioned*, and Eq. (2.16) could be solved by back-substitution,

$$x_s = \sum_{i=1}^{m+k} \left[ \frac{V_{si}}{w_i} \sum_{j=1}^m U_{ji} b_j \right]. \quad (2.18)$$

On the other hand, if some of the singular values are very small relative to others, the condition number will be large and  $\mathbf{A}$  is said to be *ill-conditioned*. In the case where the number of unknowns exceeds the number of equations by  $k$ , at least  $k$  of the singular values equal zero; the condition number is then infinite and  $\mathbf{A}$  is said to be *singular*. (Assuming the remaining  $m$  equations are linearly independent,  $k$  is said to be the *nullity* of  $\mathbf{A}$ .) In both of these cases, attempting to solve the system of equations using Eq. (2.18) would result in either a solution having large oscillations if  $\mathbf{A}$  were ill-conditioned, or division by zero if  $\mathbf{A}$  were singular.

In TSVD, singular values less than some user-defined criterion are negated (or “truncated”) by setting the corresponding  $1/w_i$  terms equal to zero. If few singular values are truncated, the radiosity and heat flux distributions usually have very irregular distributions; moreover, the solution is quite often non-physical because the radiosity and emissive power distributions assume values less than zero. As more singular values are



truncated, the solution becomes more regular. Truncating singular values also induces a residual error into the solution,  $\boldsymbol{\delta} = \mathbf{Ax} - \mathbf{b}$ ; if few singular values are truncated, however,  $\|\boldsymbol{\delta}\|_2$  is usually small compared to  $\|\mathbf{x}\|_2$ , as demonstrated in the example below. Thus, repeatedly truncating singular values leads to a feasible solution to the design problem having both a sufficiently small  $\|\mathbf{x}\|_2$  and  $\|\boldsymbol{\delta}\|_2$ , providing that such a solution exists. (This is in contrast with inverse measurement problems, where a feasible solution is known to exist.)

## 2.4 Optimization Design Methodology

Unlike the inverse methodology, the optimization design methodology solves the inverse design problem in its implicit (well-posed) form using specialized algorithms that limit the number of iterations needed to find a viable solution to the design problem. Design optimization is carried out by minimizing an objective function,  $F(\Phi)$ , which in turn is done by varying a set of design parameters contained in vector  $\Phi$  that specify the design configuration. The objective function is defined so that the optimum design configuration corresponds to the set of design parameters  $\Phi^*$  that minimizes  $F(\Phi)$ , i.e.  $F(\Phi^*) = \text{Min}[F(\Phi)]$ . In this application, the objective function is the variance of the heat flux corresponding to a particular set of design parameters from the desired heat flux distribution, evaluated at  $N_{DS}$  locations over the design surface,

$$F(\Phi) = \frac{1}{N_{DS}} \sum_{j=1}^{N_{DS}} [q_{sj}(\Phi) - q_{sj}^{target}]^2, \quad (2.19)$$

while the design parameters in  $\Phi$  are the heater settings that control the heat flux distribution over the heater surface. The heater settings that result in a heat flux distribution over the design surface that most closely matches the desired distribution corresponds with the design parameters contained within  $\Phi^*$ , which in turn is found by minimizing  $F(\Phi)$ .

It is also often desirable to impose a set of design constraints by restricting the domain of  $\Phi$ . In this application, these constraints ensure the heat flux distribution calculated over the heater surface is greater than zero but less than the maximum heating capacity of the heaters, and are enforced by satisfying inequalities of the form

$$c_i(\Phi) \leq 0, i = 1 \dots m. \quad (2.20)$$

### 2.4.1 Optimization Strategy

Many different techniques have been developed for minimizing the objective function. Gradient-based methods are most often used when the objective function is continuously differentiable and the feasible region has few local minima. These methods find  $\Phi^*$  iteratively, using the following steps. At the  $k^{th}$  iteration:

1. The design parameters are checked to see if  $\Phi^k = \Phi^*$ , which is usually done by seeing if  $\|\nabla F(\Phi^k)\|_2 < \varepsilon_{crit}$ . If  $\Phi^*$  has been identified, the process stops.
2. If  $\Phi^*$  has not been reached, a search direction,  $p^k$ , is chosen based on the curvature of  $F(\Phi^k)$ .
3. A step size,  $\alpha^k$ , is chosen. This is often done by minimizing  $f(\alpha^k) = F(\Phi^k + \alpha^k p^k)$  using a univariate minimization routine such as Newton-Raphson, bisection, or golden section searches.
4. The design parameters are updated by taking a “step” in the  $p^k$  direction,  

$$\Phi^{k+1} = \Phi^k + \alpha^k p^k.$$

Gradient-based methods differ on how the search direction is chosen. Almost all methods rely on the first-order sensitivities contained in the gradient vector,

$$\nabla F(\Phi^k) = \left\{ \frac{\partial F(\Phi^k)}{\partial \Phi_1}, \frac{\partial F(\Phi^k)}{\partial \Phi_2}, \dots, \frac{\partial F(\Phi^k)}{\partial \Phi_N} \right\}^T, \quad (2.21)$$

and some also use the second-order sensitivities contained within the Hessian matrix,

$$\nabla^2 F(\Phi^k) = \begin{bmatrix} \frac{\partial^2 F}{\partial \Phi_1^2} & \frac{\partial^2 F}{\partial \Phi_1 \partial \Phi_2} & \dots & \frac{\partial^2 F}{\partial \Phi_1 \partial \Phi_n} \\ \frac{\partial^2 F}{\partial \Phi_2 \partial \Phi_1} & \frac{\partial^2 F}{\partial \Phi_2^2} & & \\ \vdots & & \ddots & \\ \frac{\partial^2 F}{\partial \Phi_n \partial \Phi_1} & & & \frac{\partial^2 F}{\partial \Phi_n^2} \end{bmatrix}. \quad (2.22)$$

Usually, Newton's method requires the fewest iterations to find a local minimum. This is because both the first- and second-order objective function sensitivities are used to calculate the search direction,

$$\nabla^2 F(\Phi^k) \mathbf{p}^k = -\nabla F(\Phi^k). \quad (2.23)$$

It is derived based on a second-order Taylor-series expansion of  $F(\Phi)$ , and minimizes a quadratic objective function in exactly one step. (A detailed derivation is presented in Appendix A.) Although Newton's method usually requires fewer iterations to identify  $\Phi^*$  than those that only use  $F(\Phi)$  and  $\nabla F(\Phi)$  to calculate  $\mathbf{p}^k$ , it is not always more efficient because of the extra computational effort required to identify  $\nabla^2 F(\Phi)$  at each iteration. Accordingly, Newton's method should only be used when the second-order objective function sensitivities can be calculated in an efficient manner.

Besides estimating the direction of  $\Phi^*$  from  $\Phi^k$ , the magnitude of Newton's direction is also usually an accurate estimator of  $|\Phi^* - \Phi^k|$  in the vicinity of the local minimum. Accordingly, in Newton's method it is common to first assume a step size  $\alpha^k = 1$  and then to check to see if the Armijo criterion is satisfied (Bertsekas, 1999),

$$F(\Phi^k) - F(\Phi^k + \mathbf{p}^k \alpha^k) = -\mu \alpha^k \mathbf{p}^{kT} \nabla F(\Phi^k), \quad 0 < \mu < 1, \quad (2.24)$$

where  $\mu$  is a user-defined parameter. (Bertsekas, 1999a, recommends  $10^{-5} \leq \mu \leq 10^{-1}$ .)

The Armijo condition ensures that the quadratic model of  $F(\Phi)$  used to estimate  $\mathbf{p}^k$  is accurate within the vicinity of  $\Phi^k + \mathbf{p}^k \alpha^k$ , and works by checking whether the objective function decrease realized by taking a step  $\alpha^k$  in the  $\mathbf{p}^k$  direction,  $|F(\Phi^k) - F(\Phi^k + \mathbf{p}^k \alpha^k)|$ , is at least as large as some multiple of the decrease predicted by the quadratic model,  $-\mu \alpha^k \mathbf{p}^{kT} \nabla F(\Phi^k)$ . If the decrease is not large enough, then  $\alpha^k$  is reduced (usually by an order of magnitude each time) until Eq. (2.24) is satisfied.

The first step in calculating  $\nabla F(\Phi^k)$  and  $\nabla^2 F(\Phi^k)$  is to rewrite the objective function so that it is dependant on both the design parameters as well as an intermediate “system response” variable, which in turn is a function of the design parameters. In this application, the objective function is rewritten as

$$F(\Phi) = F[\Phi, \mathbf{q}_s(\Phi)], \quad (2.25)$$

where the elements of the system response,  $q_{sj}(\Phi)$ , are heat flux terms evaluated at discrete locations over the design surface. The first- and second-order objective function sensitivities are found by differentiating Eq. (2.19) with respect to the design parameters.

The terms in  $\nabla F(\Phi^k)$  are given by

$$\nabla F_p(\Phi) = \frac{\partial F(\Phi)}{\partial \Phi_p} = \frac{2}{N_{DS}} \sum_{j=1}^{N_{DS}} [q_{sj}(\Phi) - q_{sj}^{target}] \frac{\partial q_{sj}(\Phi)}{\partial \Phi_p}, \quad (2.26)$$

while the terms in  $\nabla^2 F(\Phi^k)$  are equal to

$$\nabla^2 F_{pq}(\Phi) = \frac{\partial^2 F(\Phi)}{\partial \Phi_p \partial \Phi_q} = \frac{2}{N_{DS}} \sum_{j=1}^{N_{DS}} \left\{ \frac{\partial q_{sj}(\Phi)}{\partial \Phi_p} \frac{\partial q_{sj}(\Phi)}{\partial \Phi_q} + [q_{sj}(\Phi) - q_{sj}^{target}] \frac{\partial^2 q_{sj}(\Phi)}{\partial \Phi_p \partial \Phi_q} \right\}. \quad (2.27)$$

Thus, the design parameters that produces a heat flux distribution over the design surface that most closely matches the desired distribution are found by repeatedly following the steps described above until  $F(\Phi)$  is minimized. It is important to note that the optimal heat flux distribution usually does not match the desired distribution exactly, and accordingly  $F(\Phi^*)$  is generally small but greater than zero. If the design problem has no feasible solution,  $F(\Phi^*)$  will be considerably larger than zero.

### 2.4.2 Calculation of the Design Sensitivities

Clearly, the most difficult part of calculating Eqs. (2.26) and (2.27) lies in finding the heat flux sensitivities. These are calculated using the radiosity sensitivities, which in turn are found by first directly differentiating the governing radiosity equation and then solving the resulting integral equations numerically using the procedure described in Section 2.2.

The governing radiosity equation is rewritten here to demonstrate the functional dependence of the radiosity distribution on the design parameters. Assuming that the specified heat flux and temperature distributions are controlled by the design parameters (including the heat flux distribution over the heater surface), and that the surface emissivity and enclosure geometry are independent of the design parameters, the radiosity distribution is governed by

$$q_o(u, \Phi) = b(u, \Phi) + g(u) \int_a^b q_o(u', \Phi) k(u, u') du'. \quad (2.28)$$

Differentiating Eq. (2.28) with respect to a design parameter and applying Liebnitz's rule to commute the derivative into the integral (note that the bounds of integration are constant with respect to the design parameters) results in another Fredholm integral of the second-kind,

$$\frac{\partial q_o(u, \Phi)}{\partial \Phi_p} = \frac{\partial b(u, \Phi)}{\partial \Phi_p} + g(u) \int_a^b \frac{\partial q_o(u', \Phi)}{\partial \Phi_p} k(u, u') du', \quad (2.29)$$

governing the first-order radiosity sensitivities. Likewise, differentiating Eq. (2.29) with respect to another design parameter produces another Fredholm integral equation of the second-kind,

$$\frac{\partial^2 q_o(u, \Phi)}{\partial \Phi_p \partial \Phi_q} = \frac{\partial^2 b(u, \Phi)}{\partial \Phi_p \partial \Phi_q} + g(u) \int_a^b \frac{\partial^2 q_o(u', \Phi)}{\partial \Phi_p \partial \Phi_q} k(u, u') du', \quad (2.30)$$

which governs the second-order radiosity sensitivities.

By following steps analogous to those used to solve Eq. (2.2), it can be shown that the first- and second-order radiosity sensitivities are found by solving

$$\mathbf{A} \mathbf{x}' = \mathbf{b}', \quad (2.31)$$

and

$$\mathbf{A} \mathbf{x}'' = \mathbf{b}'', \quad (2.32)$$

where  $\mathbf{A}$  is the matrix defined in Eq. (2.8),  $x_i' = \partial q_{oi}(\Phi)/\partial \Phi_p$ ,  $b_i' = \partial b_i(\Phi)/\partial \Phi_p$ ,  $x_i'' = \partial^2 q_{oi}(\Phi)/\partial \Phi_p \partial \Phi_q$  and  $b'' = \partial^2 b_i(\Phi)/\partial \Phi_p \partial \Phi_q$ . If  $\mathbf{A}$  has already been inverted or otherwise decomposed to solve for the radiosity distribution, Eqs. (2.31) and (2.32) can be easily solved by post-processing.

Once the radiosity sensitivities are known, the heat flux sensitivities are found by substituting the radiosity sensitivities into the partial derivatives of Eq. (2.10). The first-order heat flux sensitivities are found by

$$\frac{\partial q_{si}(\Phi)}{\partial \Phi_p} = \frac{\partial q_{oi}(\Phi)}{\partial \Phi_p} - \sum_{j=1}^N \frac{\partial q_{oj}(\Phi)}{\partial \Phi_p} \beta_{ij} dF_{i-stripj}, \quad (2.33)$$

while the second-order heat flux sensitivities are found by

$$\frac{\partial^2 q_{si}(\Phi)}{\partial \Phi_p \partial \Phi_q} = \frac{\partial^2 q_{oi}(\Phi)}{\partial \Phi_p \partial \Phi_q} - \sum_{j=1}^N \frac{\partial^2 q_{oj}(\Phi)}{\partial \Phi_p \partial \Phi_q} \beta_{ij} dF_{i-stripj}. \quad (2.34)$$

Finally, the heat flux sensitivities are substituted into Eqs. (2.26) and (2.27) to find  $\nabla F(\Phi)$  and  $\nabla^2 F(\Phi)$ .

## 2.5 Demonstration of Inverse and Optimization Design Methodologies

The inverse and optimization design methodologies are demonstrated and compared by using them to solve the radiant enclosure design problem shown in Fig. 2.4. The enclosure consists of a design surface having an emissivity  $\varepsilon_{DS} = 0.5$ , two adiabatic surfaces with emissivities  $\varepsilon = 1$ , and a heater surface with emissivity  $\varepsilon_{HS} = 0.9$ . The heater surface contains 24 uniformly spaced heaters, and the heat flux distribution over each heater is assumed to be uniform. Due to symmetry, the heat flux distribution over the heater surface is fully characterized using 12 design parameters,  $\Phi_h$ ,  $h = 1 \dots 12$ , where  $\Phi_h$  is equal to the heat flux over the  $h^{th}$  heater in units of  $W/m^2$ . The goal of the analysis is to find a combination of heater settings that will produce a uniform heat flux of  $q_{sDS}^{target} = -2 W/m^2$  and an emissive power  $E_{bDS} = 1 W/m^2$  over the design surface.

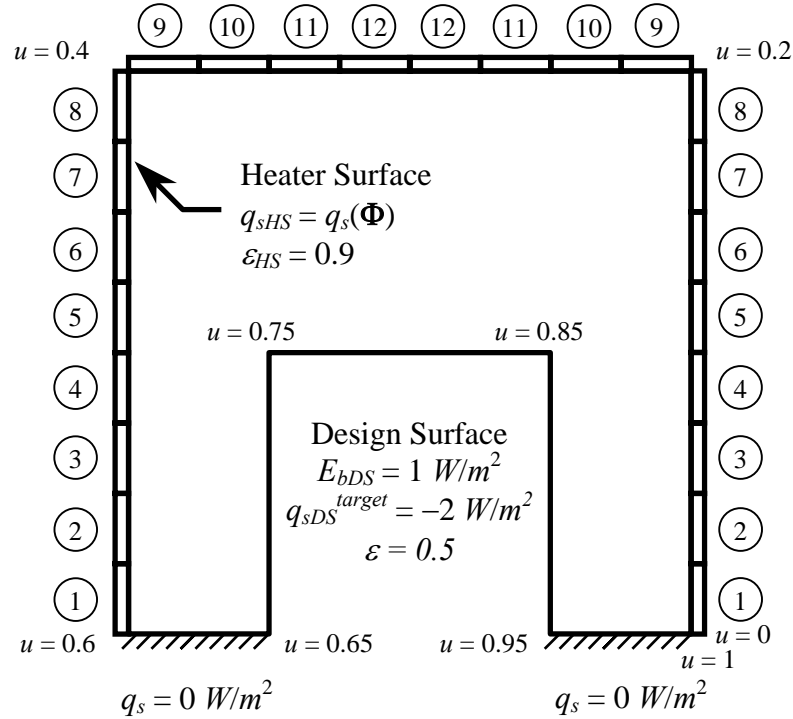


Figure 2.4: Example radiant enclosure design problem.



A total of 640 surface elements is used in both the inverse and optimization design methodologies; 384 of those elements lie on the heater surface, 192 lie on the design surface, and the remaining 64 elements are located on the two adiabatic surfaces. In order to ensure grid independence at this level of refinement, a grid refinement study was performed with  $\Phi_h^0 = 1$ ,  $h = 1 \dots 12$ , corresponding with the initial enclosure configuration used in the optimization methodology, and is shown in Fig. 2.5. The grid refinement study was performed on the objective function defined in Eq. (2.19); the grid-independent solution was taken as the value of the objective function at the highest level of refinement,  $F_\infty(\Phi^0) \approx F_{N=5120}(\Phi^0) = 0.05732$ , and the error at the lower levels of refinement was estimated by

$$E(N, \Phi) = F_N(\Phi) - F_\infty(\Phi). \quad (2.35)$$

A grid-refinement study was also performed on the degree of energy conservation, defined by

$$\% EI = \frac{\sum_{j=1}^N q_{sj} \Delta A_j}{\sum_{j=1}^N |q_{sj}| \Delta A_j} \times 100\%, \quad (2.36)$$

which equals zero for an exact solution. Assuming a uniform radiosity distribution over each surface element is a necessary step in the discretization of the governing integral equations, but also means that energy conservation is not exactly satisfied. As  $N$  becomes large, however, the discretization error gets smaller; the energy imbalance tends to zero, and  $F(\Phi)$  approaches its grid-independent solution. Using these procedures, the discretization error in  $F(\Phi^0)$  at  $N = 640$  was 0.36%, and the energy imbalance was  $-0.22\%$ . These grid refinement studies were also performed on the solutions found using the inverse design methodology and the optimization design methodology.

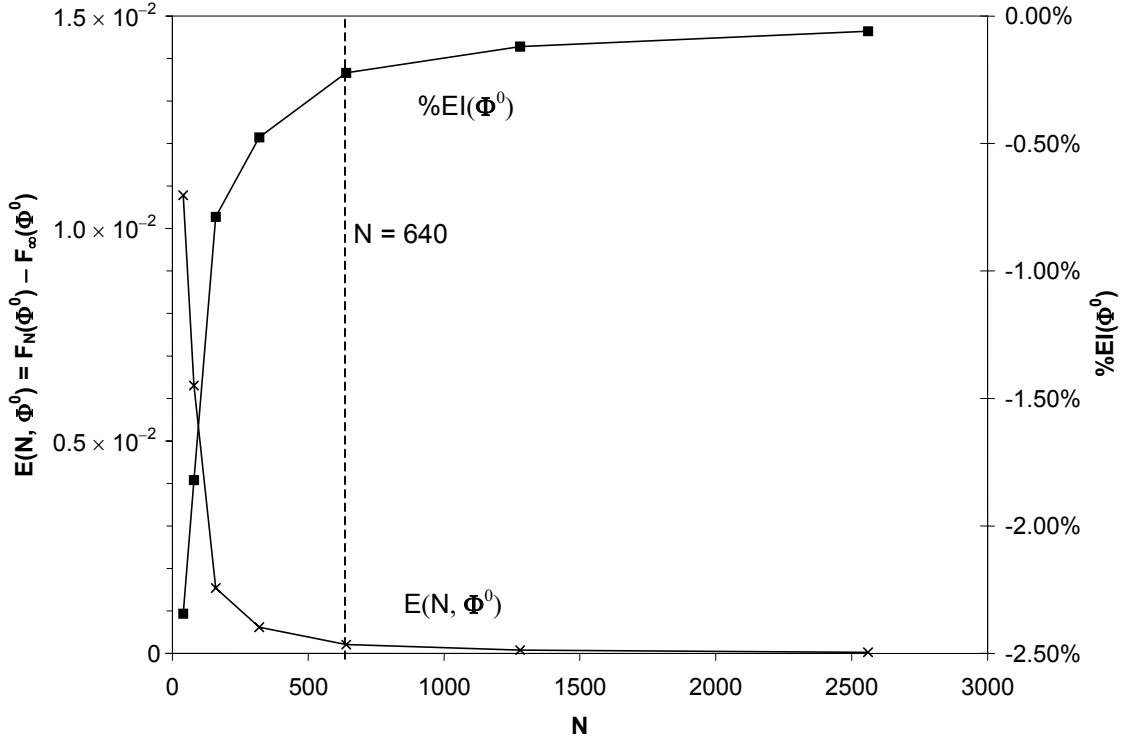


Figure 2.5: Grid refinement study for the design problem in Fig. 2.4, performed at initial configuration for optimization,  $\Phi^0$ .

### 2.5.1 Inverse Design Methodology Solution

In order to enforce a uniform heat flux over each heater, the unknown heat flux terms over the heater surface,  $q_{si}$ , were replaced by the heater settings,  $\Phi_h$ . This both reduces the degree of rank deficiency (or equivalently the nullity of  $\mathbf{A}$ ) and also acts to regularize the solution by preventing the occurrence of highly-oscillatory heat flux distributions, which are undesirable from a designer's perspective.

Although the system of linear equations used in the forward and optimization design methodologies enforces energy conservation, this is not the case in the inverse design methodology because both boundary conditions over the heater surface are

unspecified. This problem is addressed using a modification suggested by Jones (1998) in which an energy conservation equation is added to the system of equations that form the  $\mathbf{A}$  matrix,

$$\sum_{h=1}^{12} A_{heater} \Phi_h = -\frac{q_{sDS}^{target} A_{DS}}{2}, \quad (2.37)$$

where  $A_{heater}$  is the surface area of each heater and  $A_{DS}$  is the surface area of the design surface. Once the  $\mathbf{A}$  matrix is formed, it is decomposed into the three matrices of Eq. (2.17). The 460 singular values are reordered and are shown in Fig. 2.6. Eleven of these singular values equal zero due to the rank-deficiency of the  $\mathbf{A}$  matrix, which in turn is caused by the twelve unknown parameters that control the heat flux distribution over the heater surface. (Adding Eq. (2.36) to the system of equations reduces the nullity of  $\mathbf{A}$  by one.) In some other problems (Morales et al., 1996, for example), the singular values ranged over several orders of magnitude, which indicated that there were many different feasible solutions to the design problem. In contrast, however, Fig. 2.6 shows that most of the singular values are close to unity, suggesting the existence of only a few unique solutions.

The next step is to truncate the smallest singular values and to calculate the heat flux and emissive power distribution over the heater surface using Eq. (2.18). The heat flux and emissive power obtained using all the non-zero singular values ( $p = 449$ ) and the case for  $p = 448$  are shown in Figs. 2.8 and 2.9, respectively. The corresponding design parameters are included in Table 2.1, while the  $L_2$  norms of the solution and residual vectors are included in Table 2.2. Since the inverse design methodology solves the inverse design problem in its explicit form,  $T(u) = T_{DS}^{target}$  and  $q_s(u) = q_{sDS}^{target}$  for  $0.65 \leq u \leq 0.95$ .

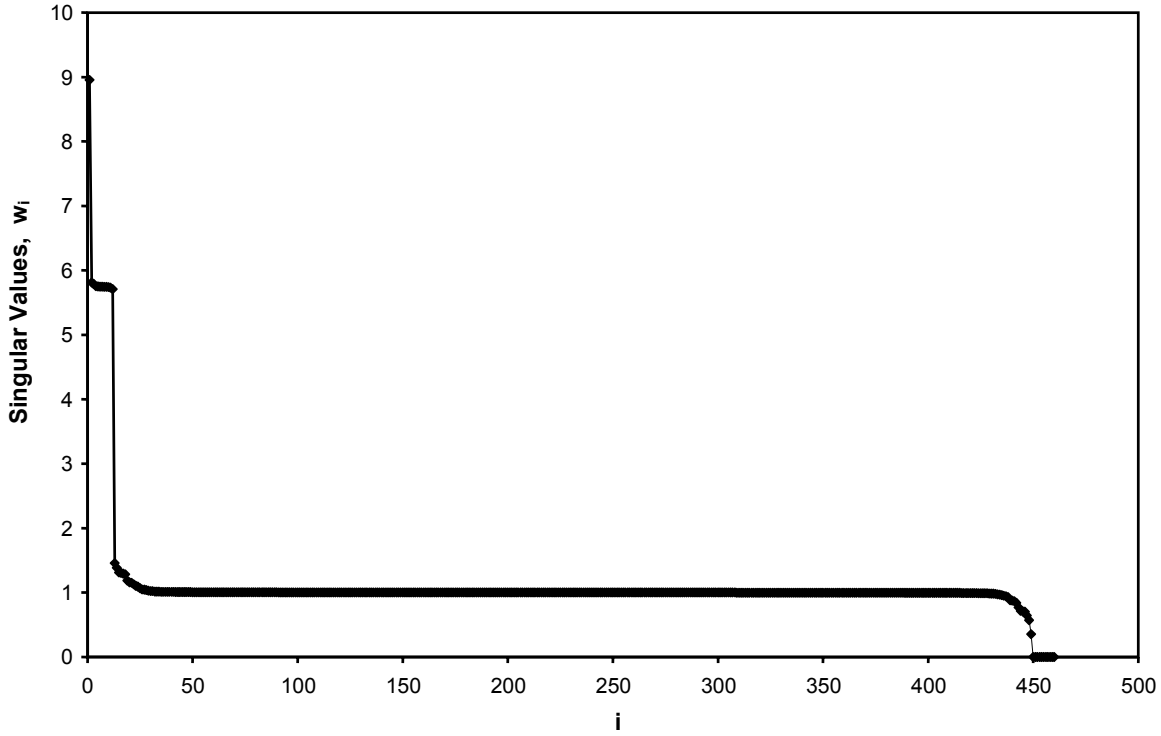


Figure 2.6: Ordered singular values found by decomposing  $\mathbf{A}$  matrix.

Since some of the singular values are equal to zero, it is not possible to generate a solution using  $p = 460$  singular values without dividing by zero in Eq. (2.18). Truncating the eleven singular values corresponding to the nullity of  $\mathbf{A}$  ( $p = 449$ ) produces a regularized solution and a very small residual, as shown in Fig. 2.8. Although the heat flux distribution can be further regularized by truncating an additional singular value ( $p = 448$ ), the solution is non-physical because  $E_b(u)$  assumes negative values over parts of the heater surface as shown in Fig. 2.8. At this point, too many singular values have been truncated and the original governing equations are no longer enforced, as shown by the large residual in Table 2.2.

Although the solution found with  $p = 449$  singular values is physically obtainable, it is not convenient to implement in a practical setting because the heat flux distribution

over some of the heaters is negative. Rather than adding heat sinks to the heater surface, the designer is more likely to simply insulate these heaters and incur the resulting deviation of the boundary conditions over the design surface. In this application, the heaters corresponding to  $\Phi_7$ ,  $\Phi_8$ ,  $\Phi_9$ , and  $\Phi_{10}$  are first set equal to zero, and the resulting heat flux distribution over the design surface (assuming that  $T(u) = T_{DS}^{target}$  over the design surface) is recalculated and plotted in Fig. 2.10, along with the solution obtained using the optimization design methodology.

	$\Phi_1$	$\Phi_2$	$\Phi_3$	$\Phi_4$	$\Phi_5$	$\Phi_6$
$p = 449$	1.8699	3.7785	3.8663	2.4447	0.9005	0.1314
$p = 448$	0.1110	-0.6260	-0.6704	-0.28903	-0.1095	0.3064
	$\Phi_7$	$\Phi_8$	$\Phi_9$	$\Phi_{10}$	$\Phi_{11}$	$\Phi_{12}$
$p = 449$	-0.5398	-0.9781	-0.0614	-0.1410	0.1364	0.5924
$p = 448$	0.4819	0.6462	0.4915	0.3746	0.2310	0.0547

Table 2.1: Heater settings found with the inverse design methodology [ $W/m^2$ ].

	$\ \mathbf{x}\ _2$	$\ \delta\ _2$
$p = 449$	$1.0297 \times 10^4$	$1.1776 \times 10^{-25}$
$p = 448$	$3.5822 \times 10^1$	$1.9291 \times 10^3$

Table 2.2:  $L_2$  norms of the solution and residual vectors corresponding to solutions found using the inverse design methodology.

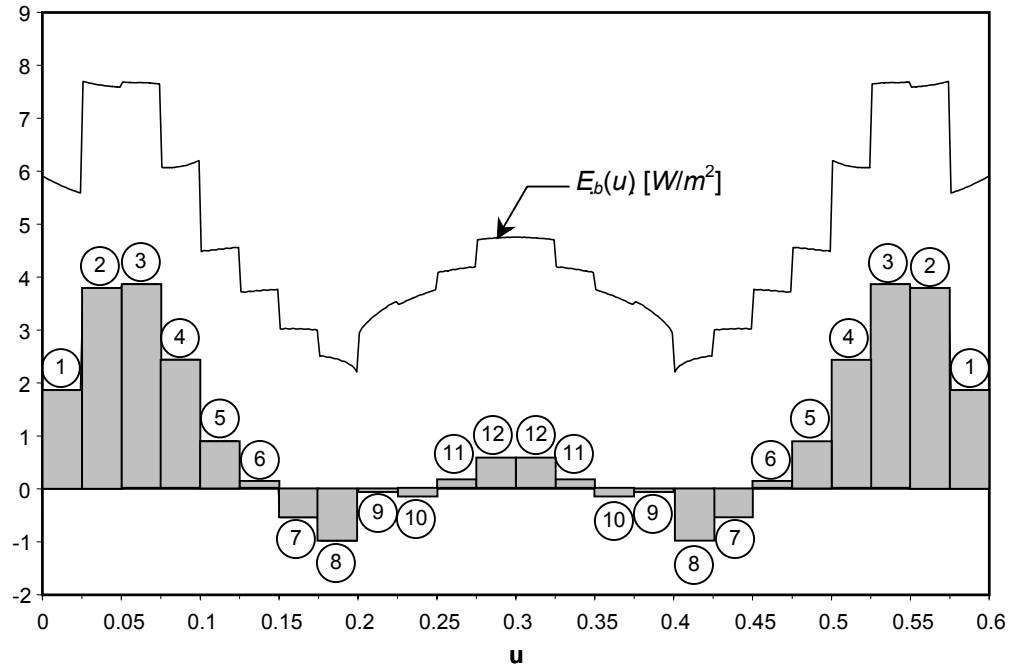


Figure 2.7: Heat flux and emissive power distributions over the heater surface found using all the non-zero singular values ( $p = 449$ ).

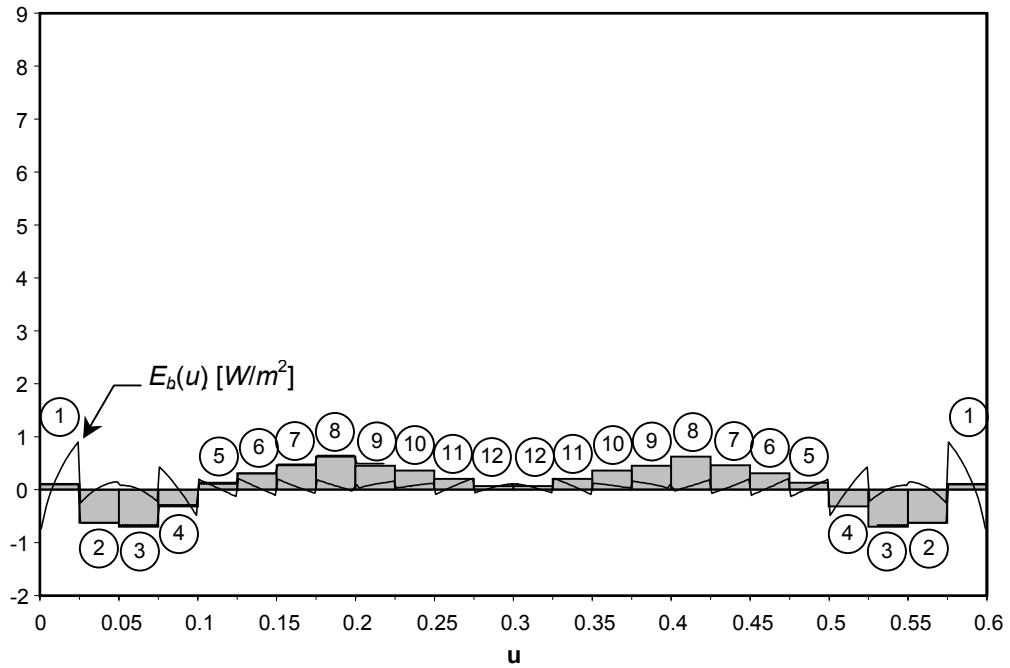


Figure 2.8: Heat flux and emissive power distributions over the heater surface found using  $p = 448$  singular values.

### 2.5.2 Optimization Design Methodology Solution

The design problem was also solved by minimizing the objective function defined in Eq. (2.19) using Newton's method. As previously stated, it is desirable from a designer's perspective to find solutions that have strictly nonnegative heat flux distributions over the design surface. Also, in a practical design setting there may also exist a maximum heater output that must not be exceeded. Solutions that satisfy both of these conditions are found by using inequality design constraints to ensure that  $l_{bh} \leq \Phi_h \leq u_{bh}$ ,  $h = 1 \dots 12$ , where  $l_{bh} = 0 \text{ W/m}^2$  and  $u_{bh}$  is arbitrarily set equal to  $10 \text{ W/m}^2$ . These constraints were enforced using an active set method, which is described in detail in Appendix A. Minimization started from  $\Phi_h = 1 \text{ W/m}^2$ ,  $h = 0 \dots 12$ , and was stopped when  $|\nabla F(\Phi)| < 10^{-4}$ .

Twelve iterations were required to identify a local minimum  $F(\Phi^*) = 3.9006 \times 10^{-4}$ , corresponding to the solution shown in Fig. 2.9 and the design parameters shown in Table 2.3. The heat flux distribution over the design surface closely matches  $q_{sDS}^{target}$ , as shown in Fig. 2.10. The maximum deviation is less than 2.1%, which is well within the tolerances demanded by most engineering applications. Another favorable property of this solution is that, due to the action of the active set method, only eight of the original twenty-four heaters are required.

$\Phi_1$	$\Phi_2$	$\Phi_3$	$\Phi_4$	$\Phi_5$	$\Phi_6$
3.1375	0	1.3909	1.8801	0	0
$\Phi_7$	$\Phi_8$	$\Phi_9$	$\Phi_{10}$	$\Phi_{11}$	$\Phi_{12}$
0	0	5.5709	0	0	0

Table 2.3: Heater settings found with the optimization design methodology [ $\text{W/m}^2$ ].

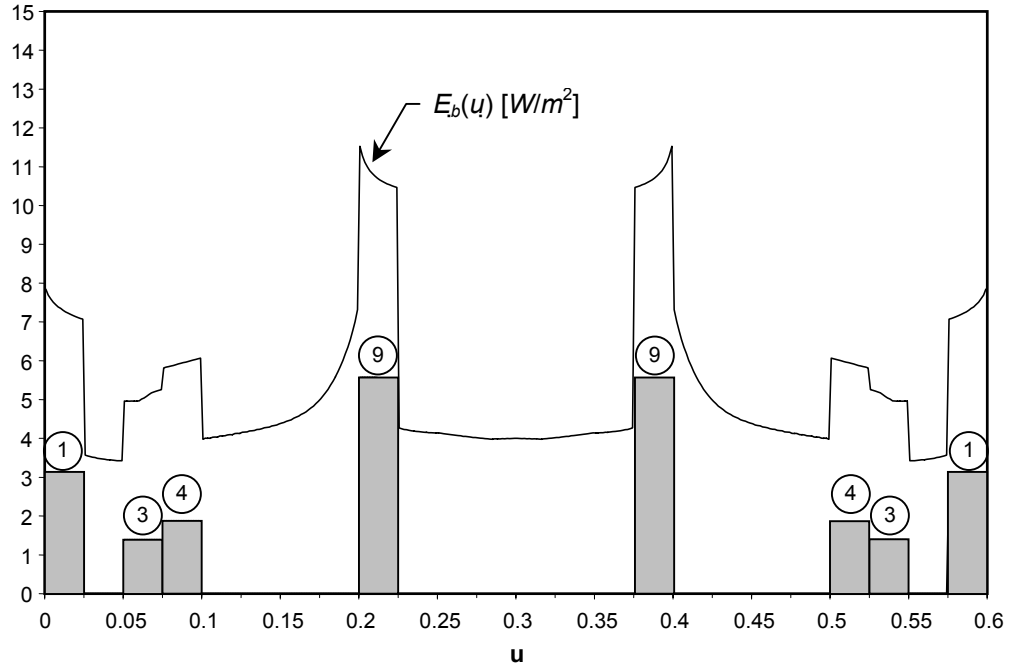


Figure 2.9: Heat flux and emissive power distributions over the heater surface found through the optimization design methodology.

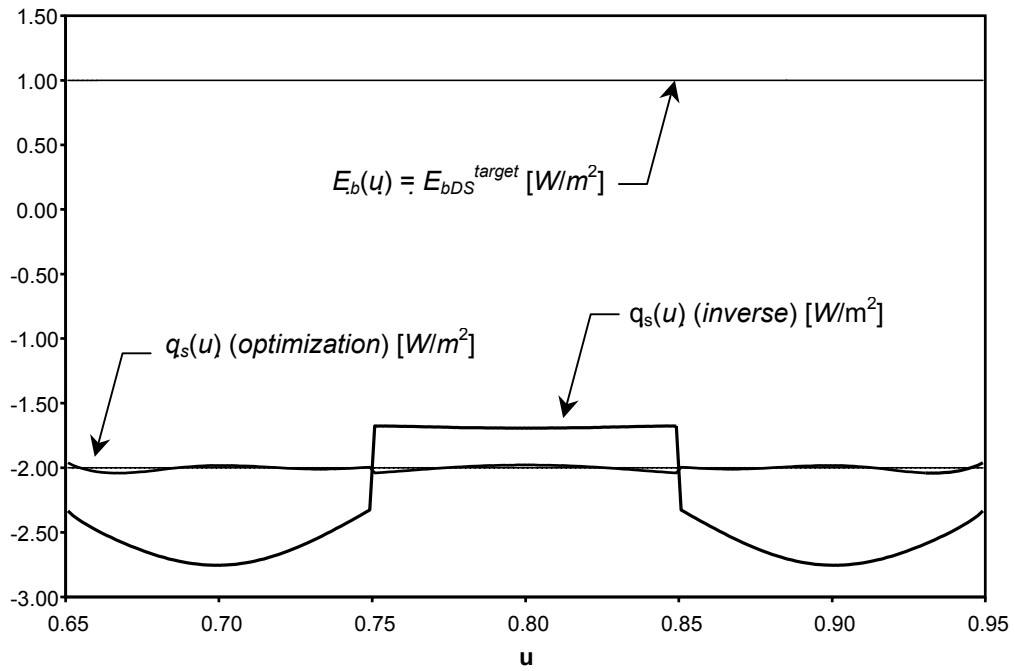


Figure 2.10: Heat flux distributions over the design surface produced by the heater settings found using the optimization design methodology and the inverse design methodology with negative heater settings set equal to zero.



### 2.5.3 Comparison of Inverse and Optimization Design Methodologies

The inverse and optimization design methodologies are compared based on their ease of formulation, required computational effort, and the quality of their final solutions.

Both methodologies identify different heater settings that produce the desired heat flux and temperature distributions over the design surface. In the solution found through the inverse design methodology, shown in Fig. 2.7, all the heaters are activated (although some of them act as heat sinks), while the optimization methodology found a solution involving only eight of the twenty-four heaters as shown in Fig. 2.9. The fact that two very different heat flux distributions over the heater surface produce nearly identical conditions over the design surface demonstrates the ill-posed nature of the problem.

The inverse design methodology is somewhat more difficult to implement, since the governing radiosity equations must be rewritten and rearranged from their natural form. In addition, the regularization techniques required to solve the ill-conditioned set of linear equations demand specialized mathematical knowledge on the part of the designer. On the other hand, the optimization design methodology solves the inverse design problem in its implicit “natural” form, and uses the same governing equations that are used in the traditional forward design methodology. The complexity of the optimization process depends largely on how the sensitivities are calculated; in this example the sensitivities are found by differentiating the governing equations, which requires some mathematical analysis. In many other cases, however, the sensitivities are estimated using a forward or central difference approximation, which doesn’t require manipulation of the governing equations. Moreover, gradient-based minimization is

more intuitive and easier to implement than the regularization procedure described in Eqs. (2.17) and (2.18).

Both the inverse and optimization design methodologies require approximately the same computational effort and storage to implement. The TSVD regularization method involves only one singular-value decomposition of the  $\mathbf{A}$  matrix, demanding approximately the same CPU time as the single  $\mathbf{A}$  matrix inversion performed during the Newton minimization. In this problem, both the inverse and optimization design methodologies solved the design problem in a matter of minutes.

The main difference between the two design methodologies lies in the quality and usefulness of the final solutions, shown in Fig. 2.10. While the solution obtained by the inverse design methodology is mathematically superior (in the sense that both boundary conditions are almost exactly satisfied over the design surface), the solution obtained through the optimization design methodology is preferable under realistic operating conditions, where it would be impractical to operate the radiant enclosure with heat sinks on the heater surface. The optimization design methodology allows for practical design considerations by enforcing constraints throughout the minimization procedure. In contrast, the inverse methodology solves the unconstrained problem through regularization, and only later can the designer modify this solution to account for the constraints. The ability of the optimization methodology to accommodate design constraints is one of the principle assets of this design technique.

## 2.6 Conclusions

This chapter presents forward, inverse, and optimization methodologies that can be used to solve for the heater settings in fixed-geometry radiant enclosure design problems. In the forward and optimization methodologies, the temperature distribution is specified over the design surface and the heater settings that produce the desired heat flux distribution are found iteratively. While the forward methodology relies solely on the designer's intuition and experience to make changes at each iteration, the optimization methodology employs a specialized minimization algorithm that both reduces the required number of iterations and greatly improves the final solution quality. In contrast, in the inverse design methodology both conditions are enforced over the design surface, while the boundary condition over the heater surface is unspecified. The resulting ill-posed problem is solved using a regularization method.

In this implementation, the governing equations for each design methodology are derived using the infinitesimal-area analysis technique. In the inverse design methodology, discretizing the governing integral equations results in a system of ill-conditioned linear equations that are solved using truncated singular value decomposition. The optimization methodology uses Newton's method to minimize the objective function, since the first- and second-order objective function sensitivities are calculated efficiently by differentiating the governing equations. Design constraints in the optimization methodology are enforced using the active set method.

The solution found using the inverse design methodology was mathematically superior to the one found through the optimization design methodology, but also involved negative heat fluxes over the heater surfaces that make it impractical to implement in an

industrial setting. The optimization design methodology avoids this by enforcing inequality design constraints throughout the optimization process to ensure that the heat flux distribution over the heater surface remains strictly nonnegative. The ability of the optimization design methodology to accommodate design constraints usually results in solutions that can be implemented in a practical setting, and is one of the main assets of this technique.

Many of the radiant enclosure design problems that arise in industry are of the type that can be solved using the design methodologies presented in this chapter; nevertheless, there are others in which the objective is to determine the optimal enclosure geometry as well as the heater settings. The optimization methodology presented here is extended to solve such problems in the next chapter.

# **Chapter 3:**

## **Geometric Optimization of Radiant Enclosures Containing Diffuse Surfaces**

### 3.1 Introduction

The optimization design methodology presented in the previous chapter is restricted to solving radiant enclosure design problems in which the goal is to find the heater settings that produce the desired heat flux and temperature distributions over the design surface. This chapter extends the optimization design methodology to solve the inverse geometry problem, where both the enclosure geometry and the heater settings can be changed in order to realize the desired conditions over the design surface.

Inverse geometry problems can broadly be separated into two types: inverse geometry identification problems, in which experimental data is used to infer the shape of a surface that could otherwise not be measured, and inverse geometry design problems where the goal is to identify the geometry that best satisfies a given design objective. Both classes of problems are usually nonlinear in nature and are therefore ill suited to the inverse and forward methodologies described in the previous chapter.

Instead, optimization methodologies have been used extensively in the non-radiative literature to solve these problems. In geometry estimation problems, the objective function is most often defined as the variance between a set of experimental measurements (most often of temperature, in thermal applications) and a set of complementary values obtained using a numerical model in which system geometry is variable. Minimization techniques are then employed to identify the geometry that minimizes the objective function. For example, Nowak et al. (2002) presented a method for estimating the shape and location of the solidus line in continuous casting applications. In this example, the solidus line is represented by a Bézier curve with its

control points set equal to the design parameters. These design parameters are modified according to a minimization routine until the objective function (defined as the variance between measured and calculated temperatures) is minimized. In another example, Huang and Chen (1999) estimated the boundary growth of a solid immersed in a moving fluid based on thermocouple measurements taken within the interior of the solid, using a similar technique.

In inverse geometry design problems, the objective function measures the “goodness” of a particular design outcome realized with that geometry. As described in the previous chapter, the optimal design is found by minimizing this objective function subject to design constraints. For example, Ebrahimi et al. (1997) determined the optimal riser geometry for an investment casting by minimizing the riser volume using a quasi-Newton algorithm, subject to constraints that ensure directional solidification from the product to the riser to prevent voids from forming in the casting. Other recent examples include Li et al. (1999), who presents an algorithm that minimizes the amount of material needed to conduct heat from a thermal source to a thermal sink, and Fabbri (1998), who determined the fin geometry that maximized heat transfer from the walls of an annular duct to a laminar fluid. Both of these cases used genetic algorithms to minimize the objective functions. (Forrest, 1993, presents a good discussion of optimization through genetic algorithms.)

Although there are many examples of geometric optimization in systems involving conduction and convection heat transfer, very little work has been carried out on problems involving thermal radiation. Most examples in the literature present

methods for optimizing the geometry of radiating fins in applications where radiation is the dominant mode of heat transfer, including Sasikumar and Balali (2002) and Krishnaprakas and Badari Narayana (2003). In these cases, the optimum design was found by performing univariate parametric studies for each design parameter, which restricts the number of design parameters that can be considered and neglects their codependence. Although such a primitive optimization strategy may be sufficient in this particular application, it is clearly unsuitable for more complex problems involving many design parameters.

This chapter presents an optimization methodology that can solve inverse design problems involving two-dimensional radiant enclosures composed of diffuse-gray walls that contain a non-participating medium, with the objective of determining the enclosure geometry and heater settings that produce the desired heat flux and temperature distribution over the design surface. The optimization is carried out using Newton's method, which requires the heat flux and the first- and second-order heat flux sensitivity distributions over the design surface to calculate the search direction. These quantities are found from the radiosity and radiosity sensitivities using a technique similar to the one presented in the previous chapter. (Geometric terms in the equations written in the previous chapter are also described in detail.) Although any type of parametric representation can be employed to represent the enclosure geometry, Bézier curves (which are a type of Non-Uniform Rational B-Spline, or NURBS curve) are especially convenient for this purpose and their use is described here. Finally, this design methodology is demonstrated by using it to solve an example design problem.



It is important to note that, unlike the technique presented in the previous chapter, this optimization methodology is restricted to the treatment of “unobstructed” enclosures, in which the enclosure interior surface is entirely concave. In these problems the kernel is guaranteed to be continuously differentiable with respect to the design parameters, which is required since the radiosity sensitivities are calculated by directly differentiating the governing integral equations.

### 3.2 Optimization Strategy

The optimization design methodology presented in this chapter solves design problems of the type shown in Fig. 3.1, where the objective is to find the enclosure geometry and heater settings that produce the desired temperature and heat flux distribution over the design surface. As in the previous chapter, this is done by first specifying the known temperature distribution over the design surface and then using the unknown heat flux distribution to define an objective function,

$$F(\Phi) = \frac{1}{N_{DS}} \sum_{j=1}^{N_{DS}} [q_{sj}(\Phi) - q_{sj}^{target}]^2. \quad (3.1)$$

The heat flux distribution is a function of the design parameters contained in  $\Phi$ , which in this case control both the enclosure geometry and heat flux distribution over the heater surface. The optimal enclosure configuration is then identified by finding the set of design parameters that minimizes  $F(\Phi)$ , i.e.  $F(\Phi^*) = \text{Min}[F(\Phi)]$ . Inequality constraints of the form

$$c_i(\Phi) \leq 0, \quad i = 1 \dots m, \quad (3.2)$$

are used to impose bounds on the design parameters throughout the minimization process, so that the enclosure size and the heat flux distribution over the heater surface lie within a specified range and also to ensure that the enclosure remains unobstructed at all times.

As will be shown in the next section, in this type of design problem the first- and second-order design sensitivities are found efficiently by post-processing the radiosity distribution. Accordingly, Newton's method is selected to minimize Eq. (3.1) coupled with an active set method to enforce Eq. (3.2). The search direction at each iteration is

solved for using the gradient vector and the Hessian matrix defined in Eqs. (2.21) and (2.22), respectively, which in turn are formed using the heat flux and first- and second-order heat flux sensitivity distributions.

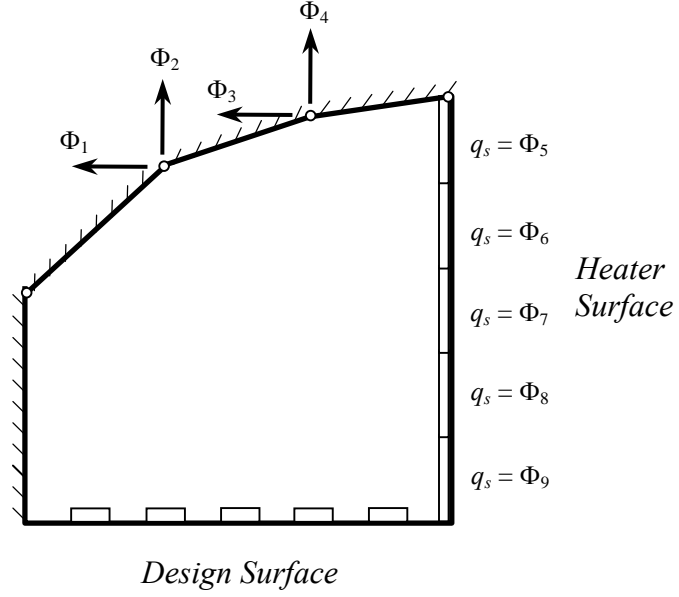


Figure 3.1: Example of a radiant enclosure inverse geometry design problem.

### 3.3 Calculation of the Design Sensitivities

The first step in calculating the design sensitivities is to represent the enclosure geometry parametrically,

$$\mathbf{r} = \mathbf{C}(u, \Phi) = \{P(u, \Phi), Q(u, \Phi)\}^T, \quad a \leq u \leq b. \quad (3.3)$$

(Note that unlike Eq. (2.1),  $\mathbf{C}(u, \Phi)$  is a function of the design parameters.) It is also assumed that the emissivity,  $\varepsilon(u)$ , is specified, and either the temperature or the heat flux is known at every point on the enclosure surface. In particular,  $T^{target}(u)$  is specified over the design surface and the heat flux distribution corresponding to the heater settings contained in  $\Phi$ ,  $q_s(u, \Phi)$ , is specified over the heater surface.

Once the enclosure is represented parametrically, it is possible to write the integral equation governing the radiosity distribution,

$$q_o(u, \Phi) = b(u, \Phi) + g(u) \int_a^b q_o(u', \Phi) k(u, u', \Phi) du', \quad (3.4)$$

where if  $u$  lies on a surface where  $T(u, \Phi)$  is specified,

$$b(u, \Phi) = \varepsilon(u) \sigma T^4(u, \Phi), \quad g(u) = 1 - \varepsilon(u), \quad (3.5)$$

or, if  $u$  lies on a surface where  $q_s(u)$  is specified,

$$b(u, \Phi) = q_s(u, \Phi), \quad g(u) = 1. \quad (3.6)$$

The kernel in Eq. (3.4),  $k(u, u', \Phi)$ , is the configuration factor from an infinitesimal area element at  $u$  to an infinitesimal area element at  $u'$  divided by  $du'$ , and is a function of the enclosure geometry. Accordingly, unlike the kernel in Eq. (2.3),  $k(u, u', \Phi)$  is dependent on the design parameters, which introduces some additional complexity into the analysis.

The kernel is given by

$$k(u, u', \Phi) = \frac{[\mathbf{n}(u, \Phi) \cdot \mathbf{e}(u, u', \Phi)] \cdot [\mathbf{n}(u', \Phi) \cdot \mathbf{e}(u, u', \Phi)]}{2 |s(u, u', \Phi)|} J(u, \Phi), \quad (3.7)$$

(Hollands, 1999) where  $\mathbf{n}(u, \Phi)$  is the surface normal vector at  $u$ ,  $\mathbf{e}(u, u', \Phi)$  is a unit vector with its tail at  $u$  pointing in the direction of  $u'$ , and  $\mathbf{s}(u, u', \Phi)$  is a vector connecting  $C(u, \Phi)$  to  $C(u', \Phi)$ ; these geometric terms are shown in Fig. 3.2. The surface discriminant,  $J(u, \Phi)$ , scales an infinitesimal increment in parametric space,  $du$ , to an infinitesimal area on the enclosure surface,  $dA(u, \Phi)$ ,

$$dA(u, \Phi) = J(u, \Phi) du. \quad (3.8)$$

All of the terms in the kernel are found directly from the parametric functions that represent the enclosure geometry:

$$\mathbf{s}(u, u', \Phi) = C(u', \Phi) - C(u, \Phi), \quad (3.9)$$

$$J(u, \Phi) = \left[ \left( \frac{\partial P(u, \Phi)}{\partial u} \right)^2 + \left( \frac{\partial Q(u, \Phi)}{\partial u} \right)^2 \right]^{1/2}, \quad (3.10)$$

$$\mathbf{e}(u, u', \Phi) = -\frac{\mathbf{s}(u, u', \Phi)}{|\mathbf{s}(u, u', \Phi)|}, \quad (3.11)$$

and

$$\mathbf{n}(u, \Phi) = \frac{1}{J(u, \Phi)} \left\{ -\frac{\partial Q(u, \Phi)}{\partial u}, \frac{\partial P(u, \Phi)}{\partial u} \right\}^T. \quad (3.12)$$

Thus, once a suitable parametric representation for the enclosure geometry has been identified, the kernel of the governing integral equation can be formed directly from  $C(u, \Phi)$  by performing algebraic operations involving differential calculus.

Once the radiosity distribution has been solved, the unknown heat flux distribution over the design surface is found by

$$q_s(u, \Phi) = q_o(u, \Phi) - \int_a^b q_o(u', \Phi) k(u, u', \Phi) du'. \quad (3.13)$$

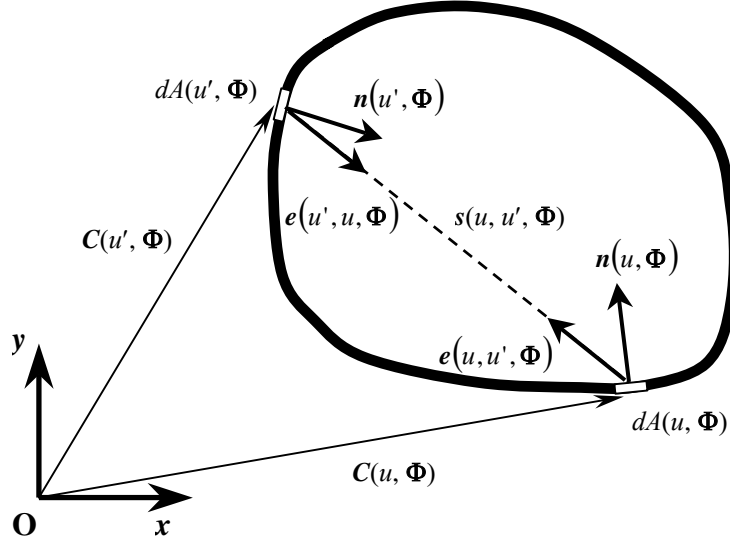


Figure 3.2: Geometric quantities used to form  $k(u, u', \Phi)$ .

The equation for the first-order radiosity sensitivity distribution is found by directly differentiating Eq. (3.3) with respect to a design parameter. In particular, the first-order radiosity sensitivity with respect to the  $p^{th}$  design parameter is governed by

$$\frac{\partial q_o(u, \Phi)}{\partial \Phi_p} = \frac{\partial b(u, \Phi)}{\partial \Phi_p} + g(u) \frac{\partial}{\partial \Phi_p} \left[ \int_a^b q_o(u', \Phi) k(u, u', \Phi) du' \right]. \quad (3.14)$$

The derivative is brought into the integral in Eq. (3.14) by applying Liebnitz's rule and noting that the integration bounds are constants,

$$\begin{aligned} \frac{\partial q_o(u, \Phi)}{\partial \Phi_p} &= \frac{\partial b(u, \Phi)}{\partial \Phi_p} + g(u) \int_a^b q_o(u', \Phi) \frac{\partial k(u, u', \Phi)}{\partial \Phi_p} du' \\ &\quad + g(u) \int_a^b \frac{\partial q_o(u', \Phi)}{\partial \Phi_p} k(u, u', \Phi) du'. \end{aligned} \quad (3.15)$$

(Since the enclosure is assumed to be unobstructed,  $k(u, u', \Phi)$  is continuous with respect to  $\Phi$  and its derivatives are known to exist.) Unlike the governing equation for the first-

order radiosity sensitivity derived in the previous chapter, Eq. (3.15) contains the radiosity as well as the first-order kernel derivatives. Nevertheless, it is still a Fredholm integral equation of the second-kind providing that  $\partial k(u, u', \Phi) / \partial \Phi_p$  has been found by differentiating Eq. (3.7) and if  $q_o(u)$  is known from solving Eq. (3.3).

By differentiating Eq. (3.15) with respect to another design parameter, say  $\Phi_q$ , and following the same algebraic steps, the second-order radiosity sensitivities are governed by

$$\begin{aligned} \frac{\partial^2 q_o(u, \Phi)}{\partial \Phi_p \partial \Phi_q} = & \frac{\partial^2 b(u, \Phi)}{\partial \Phi_p \partial \Phi_q} + g(u) \int_a^b \left[ q_o(u', \Phi) \frac{\partial^2 k(u, u', \Phi)}{\partial \Phi_p \partial \Phi_q} + \frac{\partial q_o(u', \Phi)}{\partial \Phi_p} \frac{\partial k(u, u', \Phi)}{\partial \Phi_q} \right. \\ & \left. + \frac{\partial q_o(u', \Phi)}{\partial \Phi_q} \frac{\partial k(u, u', \Phi)}{\partial \Phi_p} \right] du' + g(u) \int_a^b \frac{\partial q_o(u', \Phi)}{\partial \Phi_q} k(u, u', \Phi) du'. \end{aligned} \quad (3.16)$$

Again, Eq. (3.16) is a Fredholm integral equation of the second-kind, providing that  $q_o(u)$ ,  $\partial q_o(u) / \partial \Phi_p$ , and  $\partial q_o(u) / \partial \Phi_q$  have been found by solving Eqs. (3.4) and (3.15), and  $\partial k(u, u', \Phi) / \partial \Phi_p$ ,  $\partial k(u, u', \Phi) / \partial \Phi_q$ , and  $\partial^2 k(u, u', \Phi) / \partial \Phi_p \partial \Phi_q$  have been calculated by differentiating Eq. (3.5).

Finally, once the radiosity sensitivities have been solved, the first- and second-order heat flux sensitivities needed to calculate the gradient vector and the Hessian matrix are found by substituting the radiosity sensitivities into the derivatives of Eq. (3.12),

$$\frac{\partial q_s(u, \Phi)}{\partial \Phi_p} = \frac{\partial q_o(u, \Phi)}{\partial \Phi_p} - \int_a^b \left[ \frac{\partial q_o(u', \Phi)}{\partial \Phi_p} k(u, u', \Phi) + q_o(u', \Phi) \frac{\partial k(u, u', \Phi)}{\partial \Phi_p} \right] du', \quad (3.16)$$

and

$$\begin{aligned}
\frac{\partial^2 q_s(u, \Phi)}{\partial \Phi_p \partial \Phi_q} &= \frac{\partial^2 q_o(u, \Phi)}{\partial \Phi_p \partial \Phi_q} - \int_a^b \left[ \frac{\partial^2 q_o(u', \Phi)}{\partial \Phi_p \partial \Phi_q} k(u, u', \Phi) + \frac{\partial q_o(u', \Phi)}{\partial \Phi_p} \frac{\partial k(u, u', \Phi)}{\partial \Phi_q} \right. \\
&\quad \left. + \frac{\partial q_o(u', \Phi)}{\partial \Phi_q} \frac{\partial k(u, u', \Phi)}{\partial \Phi_p} + q_o(u', \Phi) \frac{\partial^2 k(u, u', \Phi)}{\partial \Phi_p \partial \Phi_q} \right] du', \quad (3.17)
\end{aligned}$$

respectively.



### 3.4 Solution of the Governing Equations

Since analytical solutions to Fredholm integral equations of the second kind are usually intractable, Eqs. (3.4), (3.15), and (3.16) must be solved numerically using the procedure described in the previous chapter.

First, consider the solution for the radiosity distribution. The parametric domain is divided into  $N$  elements, the  $i^{th}$  element centered at  $u_i$  with a width  $\Delta u_i$ ; each of these elements corresponds to an area element  $\Delta A_{stripi}$  on the enclosure surface that is infinitely long but has a finite width, as shown in Fig. 2.2. Next, the integrals over the entire domain of  $u$ ,  $a \leq u \leq b$ , are rewritten as the sum of  $N$  integrals, with the  $j^{th}$  integral taken over the domain  $u_j - \Delta u_j/2 \leq u \leq u_j + \Delta u_j/2$ .

Once this is done, the radiosity is assumed to be uniform over each area element, and is extracted from each of the integrals. These integrals now only contain the kernel of Eq. (3.8) and its derivatives; carrying these integrals out results in

$$q_{oi}(\Phi) = b_i(\Phi) + g_i \sum_{j=1}^N q_{oj}(\Phi) dF_{i-stripj}(\Phi), \quad (3.19)$$

where  $q_{oi}(\Phi) = q_o(u_i, \Phi)$ ,  $b_i = b(u_i, \Phi)$ ,  $g_i = g(u_i)$ , and  $dF_{i-stripj}(\Phi)$  is the infinitesimal configuration factor from  $u_i$  to the finite area element  $dA_{stripj}$  centered on  $u_j$ . Assuming that the enclosure is parameterized in a counter-clockwise manner (so that  $\hat{n}(u_i, \Phi)$  points into the enclosure),  $dF_{i-stripj}(\Phi)$  is given by

$$dF_{i-stripj}(\Phi) = \frac{1}{2} \left[ \frac{s(u_i \cdot u_j - \Delta u_j/2, \Phi) \times n(u_i, \Phi)}{|s(u_i \cdot u_j - \Delta u_j/2, \Phi)|} - \frac{s(u_i \cdot u_j - \Delta u_j/2, \Phi) \times n(u_i, \Phi)}{|s(u_i \cdot u_j - \Delta u_j/2, \Phi)|} \right], \quad (3.20)$$

(Hollands, 1999) with the terms in Eq. (3.20) shown in Fig. 3.3.

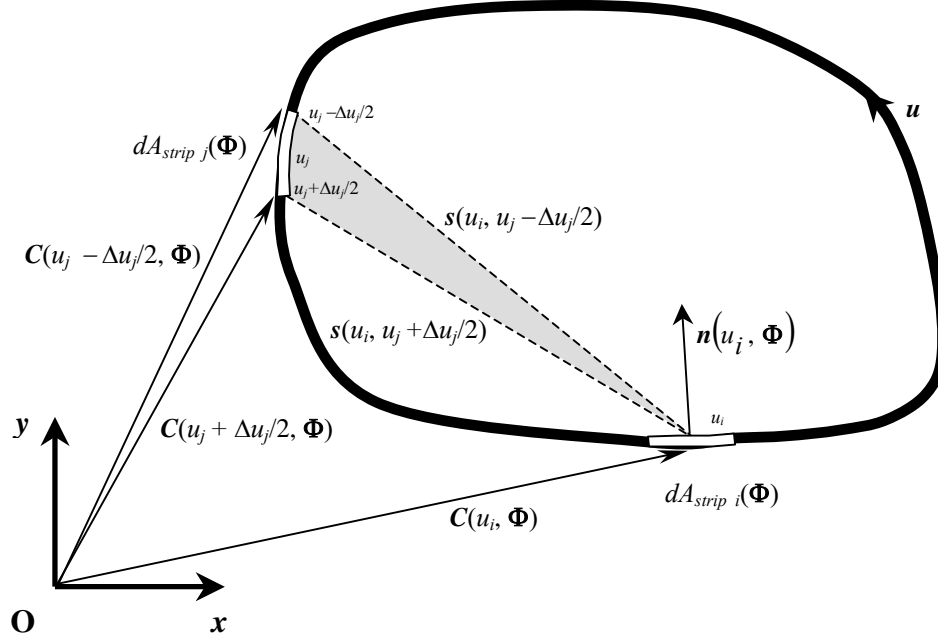


Figure 3.3: Calculation of  $dF_{i-stripj}(\Phi)$ .

Writing Eq. (3.19) for each element results in a system of  $N$  equations containing  $N$  unknowns, which can be rearranged into a matrix equation,

$$\mathbf{A}(\Phi)\mathbf{x}(\Phi)=\mathbf{b}(\Phi), \quad (3.21)$$

where  $\mathbf{x}(\Phi) = \{q_{o1}(\Phi), q_{o2}(\Phi), \dots, q_{oN}(\Phi)\}^T$ ,  $\mathbf{b}(\Phi) = \{b_1(\Phi), b_2(\Phi), \dots, b_N(\Phi)\}^T$ , and

$$A_{ij}(\Phi)=\begin{cases} 1 & i=j \\ -g_i dF_{i-stripj}(\Phi) & i \neq j \end{cases}. \quad (3.22)$$

Matrix  $\mathbf{A}$  is usually well-conditioned and can be inverted using traditional linear algebra techniques such as LU decomposition and Gauss-Seidel iteration. Finally, the unknown

heat flux values are calculated by substituting the radiosity values into the discretized form of Eq. (3.13),

$$q_{si}(\Phi) = q_{oi}(\Phi) - \sum_{j=1}^N q_{oj}(\Phi) dF_{i-stripj}(\Phi). \quad (3.23)$$

The first- and second-order radiosity sensitivities are found in a similar way. First, the domain of the integrals in Eqs. (3.15) and (3.16) are subdivided using the same parametric quadrature as before. Next, the radiosity, and first- and second-order radiosity sensitivities are assumed uniform over each finite area element and are extracted from the integrals. Carrying out the remaining integrals (containing the kernel and kernel sensitivities) results in

$$\frac{\partial q_{oi}(\Phi)}{\partial \Phi_p} = \frac{\partial b(\Phi)}{\partial \Phi_p} + g_i \sum_{j=1}^N q_{oj}(\Phi) \frac{\partial [dF_{i-stripj}(\Phi)]}{\partial \Phi_p} + g_i \sum_{j=1}^N \frac{\partial q_{oj}(\Phi)}{\partial \Phi_p} dF_{i-stripj}(\Phi), \quad (3.24)$$

for the first-order sensitivities, and

$$\begin{aligned} \frac{\partial^2 q_{oi}(\Phi)}{\partial \Phi_p \partial \Phi_q} = & \frac{\partial^2 b(\Phi)}{\partial \Phi_p \partial \Phi_q} + g_i \sum_{j=1}^N \left\{ q_{oj}(\Phi) \frac{\partial^2 [dF_{i-stripj}(\Phi)]}{\partial \Phi_p \partial \Phi_q} + \frac{\partial q_{oj}(\Phi)}{\partial \Phi_p} \frac{\partial [dF_{i-stripj}(\Phi)]}{\partial \Phi_q} \right. \\ & \left. + \frac{\partial q_{oj}(\Phi)}{\partial \Phi_q} \frac{\partial [dF_{i-stripj}(\Phi)]}{\partial \Phi_p} \right\} + g_i \sum_{j=1}^N \left[ \frac{\partial^2 q_{oj}(\Phi)}{\partial \Phi_p \partial \Phi_q} dF_{i-stripj}(\Phi) \right], \quad (3.25) \end{aligned}$$

for the second-order sensitivities.

If the radiosity distribution has been found by solving Eq. (3.21), and if  $\partial [dF_{i-stripj}(\Phi)] / \partial \Phi_p$  has been calculated by differentiating Eq. (3.20), then the first-order radiosity values are the only unknowns in Eq. (3.24). Writing Eq. (3.24) for each element results in another system of  $N$  equations containing  $N$  unknowns, which is rearranged into a matrix equation

$$\mathbf{A}(\Phi) \mathbf{x}'(\Phi) = \mathbf{b}'(\Phi), \quad (3.26)$$

where  $\mathbf{A}(\Phi)$  is defined in Eq. (3.22),  $\mathbf{x}'(\Phi) = \{\partial q_{o1}(\Phi)/\partial\Phi_p, \partial q_{o2}(\Phi)/\partial\Phi_p, \dots, \partial q_{oN}(\Phi)/\partial\Phi_p\}^T$ , and

$$b_i'(\Phi) = \frac{\partial b_i(\Phi)}{\partial\Phi_p} + g_i \sum_{j=1}^N q_{oj}(\Phi) \frac{\partial[dF_{i-stripj}(\Phi)]}{\partial\Phi_p}. \quad (3.27)$$

Likewise, once the radiosity and first-order radiosity sensitivities have been found by solving Eq. (3.21) and Eq. (3.26), and if  $\partial^2[dF_{i-stripj}(\Phi)]/\partial\Phi_p\partial\Phi_q$  has been calculated, then the second-order radiosity sensitivities are the only unknowns in Eq. (3.25). Writing Eq. (3.25) for each element results in a third system of  $N$  equations containing  $N$  unknowns, which is rearranged to form

$$\mathbf{A}(\Phi) \mathbf{x}''(\Phi) = \mathbf{b}''(\Phi), \quad (3.28)$$

where  $\mathbf{A}(\Phi)$  is again defined in Eq. (3.22),  $\mathbf{x}''(\Phi) = \{\partial^2 q_{o1}(\Phi)/\partial\Phi_p\partial\Phi_q, \partial^2 q_{o2}(\Phi)/\partial\Phi_p\partial\Phi_q, \dots, \partial^2 q_{oN}(\Phi)/\partial\Phi_p\partial\Phi_q\}^T$ , and

$$b_i''(\Phi) = \frac{\partial^2 b_i(\Phi)}{\partial\Phi_p\partial\Phi_q} + g_i \sum_{j=1}^N \left\{ q_{oj}(\Phi) \frac{\partial^2[dF_{i-stripj}(\Phi)]}{\partial\Phi_p\partial\Phi_q} + \frac{\partial q_{oj}(\Phi)}{\partial\Phi_p} \frac{\partial[dF_{i-stripj}(\Phi)]}{\partial\Phi_q} + \frac{\partial q_{oj}(\Phi)}{\partial\Phi_q} \frac{\partial[dF_{i-stripj}(\Phi)]}{\partial\Phi_p} \right\}. \quad (3.29)$$

Finally, the first- and second-order heat flux sensitivities are found by substituting the radiosity sensitivities into the discretized forms Eqs. (3.17) and (3.18),

$$\frac{\partial q_{si}(\Phi)}{\partial\Phi_p} = \frac{\partial q_{oi}(\Phi)}{\partial\Phi_p} - \sum_{j=1}^N \left\{ \frac{\partial q_{oj}(\Phi)}{\partial\Phi_p} dF_{i-stripj}(\Phi) + q_{oj}(\Phi) \frac{\partial[dF_{i-stripj}(\Phi)]}{\partial\Phi_p} \right\}, \quad (3.30)$$

and

$$\begin{aligned} \frac{\partial^2 q_{si}(\Phi)}{\partial \Phi_p \partial \Phi_q} = & \frac{\partial^2 q_{oi}(\Phi)}{\partial \Phi_p \partial \Phi_q} - \sum_{j=1}^N \left\{ \frac{\partial^2 q_{oj}(\Phi)}{\partial \Phi_p \partial \Phi_q} dF_{i-stripj}(\Phi) + \frac{\partial q_{oj}(\Phi)}{\partial \Phi_p} \frac{\partial [dF_{i-stripj}(\Phi)]}{\partial \Phi_q} \right. \\ & \left. + \frac{\partial q_{oj}(\Phi)}{\partial \Phi_q} \frac{\partial [dF_{i-stripj}(\Phi)]}{\partial \Phi_p} + q_{oj}(\Phi) \frac{\partial^2 [dF_{i-stripj}(\Phi)]}{\partial \Phi_p \partial \Phi_q} \right\}, \end{aligned} \quad (3.31)$$

respectively.

As in the optimization methodology discussed in the previous chapter, the  $\mathbf{A}(\Phi)$  matrix is common to the radiosity equation as well as the first- and second-order radiosity sensitivity equations. Furthermore, since  $\mathbf{b}'(\Phi)$  is a function of the radiosity distribution and  $\mathbf{b}''(\Phi)$  is a function of the radiosity and first-order radiosity sensitivities, the radiosity sensitivities are efficiently calculated by first decomposing  $\mathbf{A}(\Phi)$  to solve Eq (3.21), and then using this decomposition to solve Eqs. (3.26) and (3.28) sequentially. Once this is done, the heat flux and heat flux sensitivities are found using Eqs. (3.23), (3.30), and (3.31). Unlike the optimization methodology discussed in the previous chapter, however,  $\mathbf{A}(\Phi)$  is dependant on the design parameters and must be decomposed at least once during every minimization iteration.

### 3.5 Parametric Representation of Enclosure Surfaces through NURBS

Although any parametric representation can be used to represent the enclosure geometry, non-uniform rational B-spline (NURBS) curves have several unique properties that are advantageous for geometric optimization. This type of parameterization has become the standard method of curve and surface representation in CAD/CAM and computer graphics applications, and was first applied to represent radiant enclosure geometry by Daun and Hollands (2001). Detailed descriptions of NURBS curves are provided by Piegel and Tiller (1997) and Farin (1997).

In general, NURBS curves are formed by multiplying a set of points  $\{\mathbf{P}_i, i = 0 \dots n\}$ , with a corresponding set of scalar functions,  $\{N_{i,p}(u), i = 0 \dots n\}$ :

$$\mathbf{C}(u) = \sum_{i=0}^n \mathbf{P}_i N_{i,p}(u), \quad 0 \leq u \leq 1. \quad (3.32)$$

(It is customary to parameterize so that  $0 \leq u \leq 1$ , but this can be adjusted as necessary.)

Each point is called a *control point*, and the set of control points forms a *control polygon*.

The scalar functions are *basis functions*, which are ratios of two polynomials of order  $p$  defined in a piecewise manner over the domain of  $u$ . The present application uses Bézier curves to represent the enclosure geometry, which are a special subclass of NURBS curves where the polynomial in the basis function denominator is equal to unity and the number of control points exceeds the order of the polynomial in the numerator by exactly one. An example of a Bézier curve is shown in Fig. 3.4, and the corresponding basis functions are plotted in Fig. 3.5.

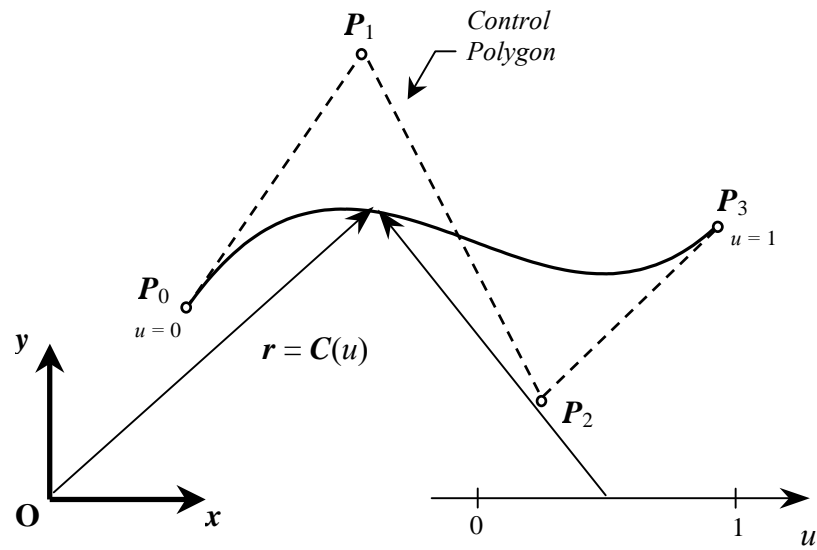


Figure 3.4: Example of a Bézier curve.

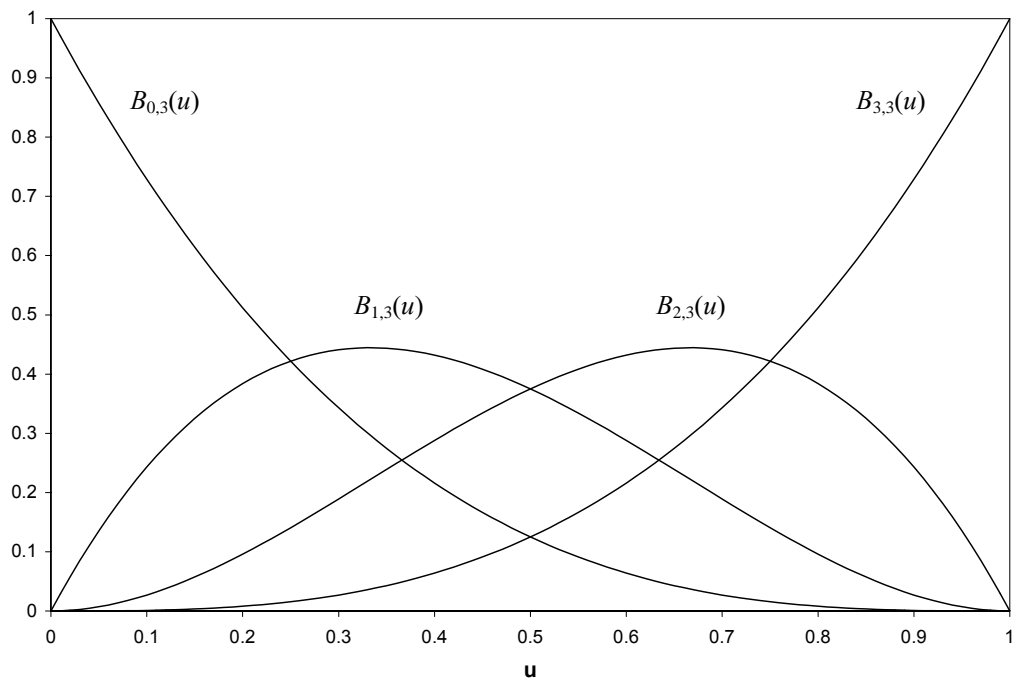


Figure 3.5: Basis functions corresponding to the Bézier curve in Fig.3.4.

One of the main advantages of parametric representation through NURBS is the computational efficiency with which points on the curve are calculated. This arises from the recursive generation of the polynomials in the NURBS basis function, using the next lower-order polynomials. For example, while Bézier basis functions are defined explicitly by

$$B_{i,p}(u) = \frac{p!}{i!(p-i)!} u^i (1-u)^{p-i}, \quad 0 \leq u \leq 1, \quad (3.33)$$

they can also be formed recursively using

$$B_{i,p}(u) = (1-u)B_{i-1,p-1}(u) + uB_{i,p-1}(u), \quad 0 \leq u \leq 1, \quad (3.34)$$

with  $B_{-1,p-1}(u) = B_{p,p-1}(u) = 0$  and  $B_{0,0}(u) = 1$ , as shown in Fig. 3.6.

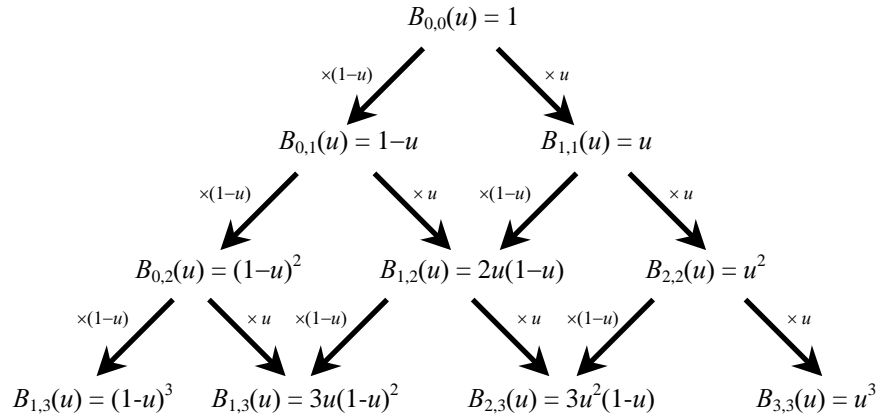


Figure 3.6: Recursive formulation of cubic Bézier basis functions.

Another favorable property of NURBS curves is the availability of the curve derivatives, which are used in this application to calculate the normal vector and surface discriminant defined in Eqs. (3.10) and (3.12), respectively. The curve derivatives are



found by differentiating Eq. (3.32) with respect to  $u$  and noting that the control points are independent of  $u$ :

$$\frac{\partial C(u)}{\partial u} = \sum_{i=0}^n P_i \frac{\partial N_{i,p}(u)}{\partial u}. \quad (3.35)$$

The recursive formulation of the basis functions permits the efficient calculation of the analytical derivatives through numerical routines. For example, the derivatives of Bézier basis functions are given by,

$$\frac{\partial B_{i,p}(u)}{\partial u} = p [B_{i-1,p-1}(u) - B_{i,p-1}(u)], \quad (3.36)$$

with  $B_{-1,p-1}(u) = B_{p,p-1}(u) = 0$ .

In the context of geometric optimization, parametric representation through NURBS presents further advantages. The simplest way of modifying the enclosure geometry is to set the design parameters equal to the coordinates of selected control points; since the basis functions are independent of the control points, they do not need to be recalculated every time a design parameter is adjusted. Also, because all the basis functions  $\{N_{i,p}(u), i = 0 \dots n\}$  are of the same order of magnitude, the curve geometry is equally sensitive to the location of each control point. This helps ensure that the objective function sensitivities have approximately the same magnitude, which improves the scaling of the optimization problem. (Objective function scaling is discussed further in Appendix A.)

Finally, the location and shape of the control points can be used to predict the location and shape of the curve. In particular, the curve begins at the first control point,  $C(0) = P_0$ , ends at the last control point,  $C(1) = P_n$ , and is always contained within the

convex hull of the control polygon. This is a consequence of the partition-of-unity property of NURBS basis functions,

$$\sum_{i=0}^n N_{i,p}(u) = 1. \quad (3.37)$$

and the fact that  $N_{0,p}(0) = N_{n,p}(1) = 1$ , as shown in Fig. 3.5. Thus, the domain of the curve can be restricted by constraining the size and shape of the control polygon. This property is later used to restrict the height and width of the enclosure and to ensure that the enclosure remains unobstructed throughout the optimization process.

### 3.6 Implementation and Demonstration

The optimization methodology described in the above sections is demonstrated by applying it to design the two-dimensional enclosure shown in Fig. 3.7. The enclosure consists of four sides; the design surface is located at the bottom of the enclosure and has a width  $W = 1 \text{ m}$ , while the two vertical surfaces are adiabatic with heights  $H = 0.5 \text{ m}$ . The top surface is represented by a cubic Bézier curve; it is heated over the left- and right-hand sides and is insulated in the middle. The design surface has an emissivity  $\varepsilon_{DS} = 0.5$ , the adiabatic surfaces (including both vertical surfaces and the insulated portion of the top surface) have emissivities of  $\varepsilon_{RS} = 0.7$ , and the emissivity of the heated portion of the top surface is  $\varepsilon_{HS} = 0.9$ . The shape of the upper surface is governed by two design parameters,  $\Phi = \{\Phi_1, \Phi_2\}^T$ , which control the  $x$ - and  $y$ - coordinates of the middle two control points of the Bézier curve, respectively. The heat flux distributions over the heated regions of the top surface are uniform and are given by

$$q_{sHS}(\Phi) = \frac{1}{L_{HS}(\Phi)}, \quad (3.38)$$

where  $L_{HS}(\Phi)$  is the chord length of the heated regions of the top surface, found by

$$L_{HS}(\Phi) = 2 \int_{0.25}^{0.3125} J(u) du. \quad (3.39)$$

Thus, while the heat flux over the heated regions of the top surface is a function of the design parameters, the total amount of heat entering the enclosure through the heaters (and leaving through the design surface) is independent of the design parameters.



The upper and lower bounds on  $\Phi_1$  and the lower bound on  $\Phi_2$  ensure that the enclosure geometry remains unobstructed, while the upper bound on  $\Phi_2$  defines the maximum enclosure height. The feasible region is plotted in Fig. 3.8.

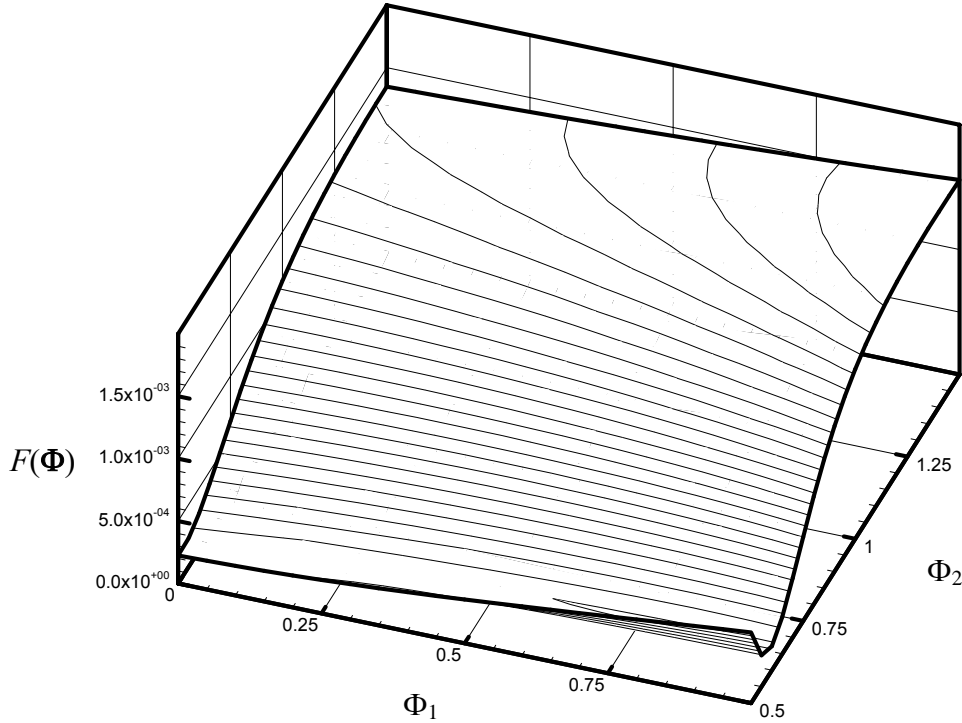


Figure 3.8: Plot of  $F(\Phi)$  over the feasible region.

Two different minimizations were carried out; the first from  $\Phi^{0,1} = \{0.8, 1.3\}^T$ , and the second starting at  $\Phi^{0,2} = \{0.5, 1\}^T$ . The optimization was carried out using Newton's method, and the constraints in Eqs. (3.40) and (3.41) were enforced using the active set method. A total of 512 elements was used to calculate the heat flux distribution; grid refinement studies based on the objective function error and the energy imbalance, defined in Eqs. (2.35) and (2.36) respectively, were performed using the

initial set of design parameters to ensure this level number of elements was sufficient. In each case, the grid-independent value of  $F(\Phi)$  was taken at the highest level of grid refinement,  $F_\infty(\Phi^0) \approx F_{N=2048}(\Phi^0)$ , and the results are summarized in Fig. 3.9 and Table 3.1. Similar studies were carried out at the local minima,  $\Phi^{*,1}$  and  $\Phi^{*,2}$ , in order to ensure grid independence.

	$F_\infty(\Phi)$	$F_{N=512}(\Phi)$	$E(N=512, \Phi)$	$\%EI(\Phi)$
$\Phi^{0,1}$	$1.497 \times 10^{-3}$	$1.495 \times 10^{-3}$	-0.16 %	0.112 %
$\Phi^{0,2}$	$1.173 \times 10^{-3}$	$1.171 \times 10^{-3}$	-0.18 %	0.021 %

Table 3.1: Grid refinement statistics for initial enclosure configurations shown in Figs. 3.11 (a) and 3.13 (a).

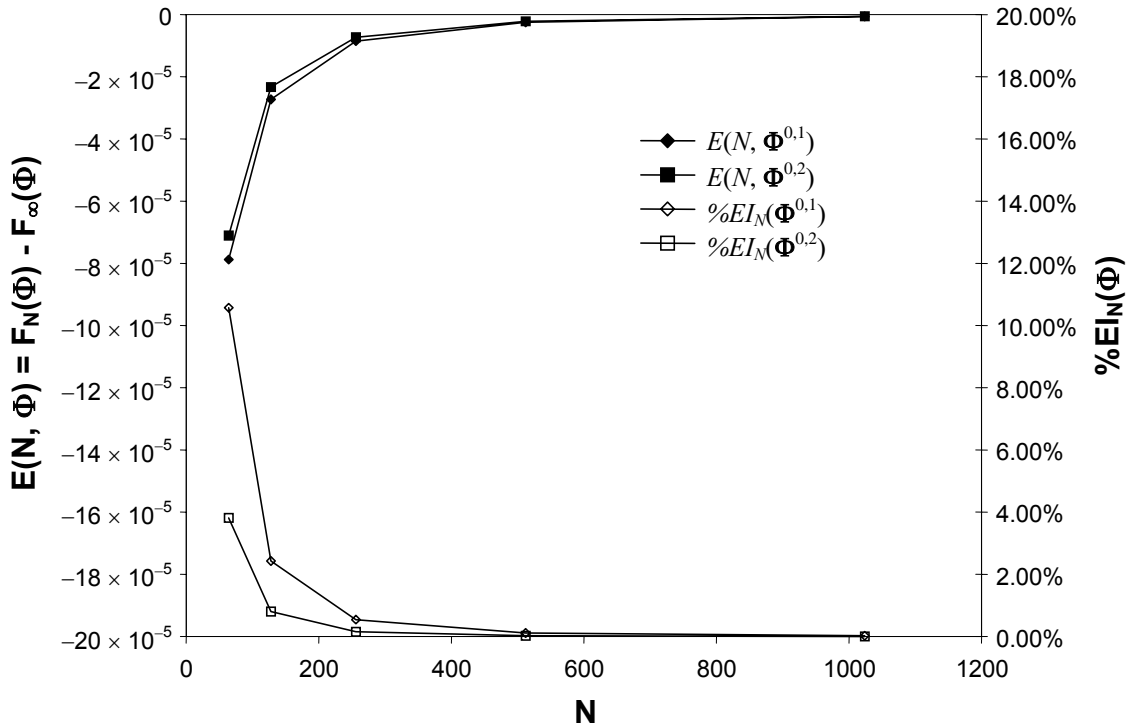


Figure 3.9: Grid refinement studies for the design problem shown in Fig. 3.7.

The first optimization started at  $\Phi^{0,1} = \{0.8, 1.3\}^T$  and required four steps to identify a local minimum at  $\Phi^{*,1} = \{1, 0.5896\}^T$ , corresponding to  $F(\Phi^{*,1}) = 1.915 \times 10^{-4}$ . The feasible region and minimization path is plotted in Fig. 3.10, and the initial and final enclosure configurations and heat flux distributions over the design surface are shown in Figs. 3.11 and 3.14, respectively.

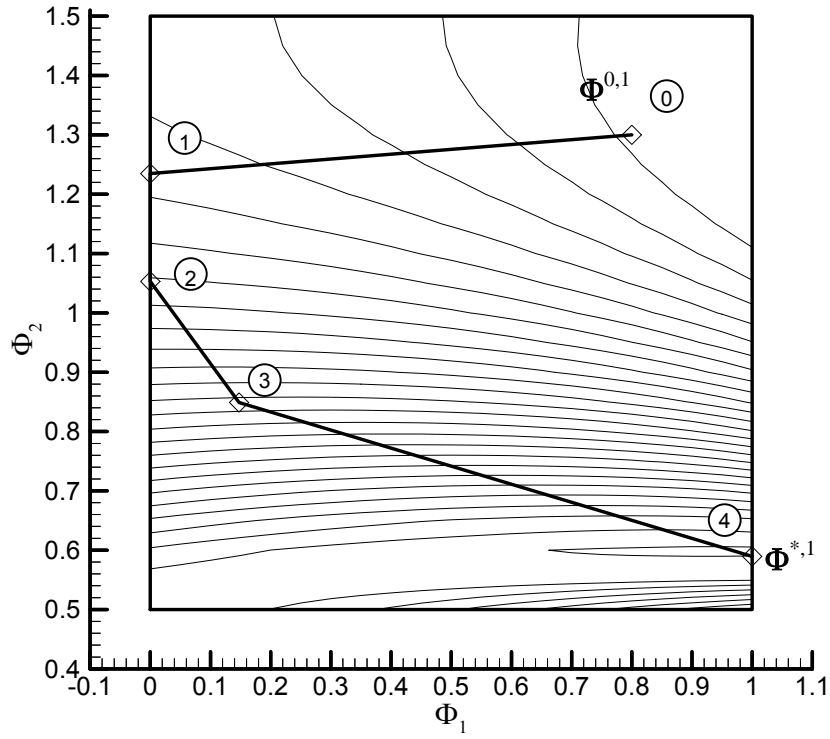


Figure 3.10: Minimization path starting from  $\Phi^{0,1} = \{0.8, 1.3\}^T$  and ending at  $\Phi^{*,1} = \{1, 0.5896\}^T$ .

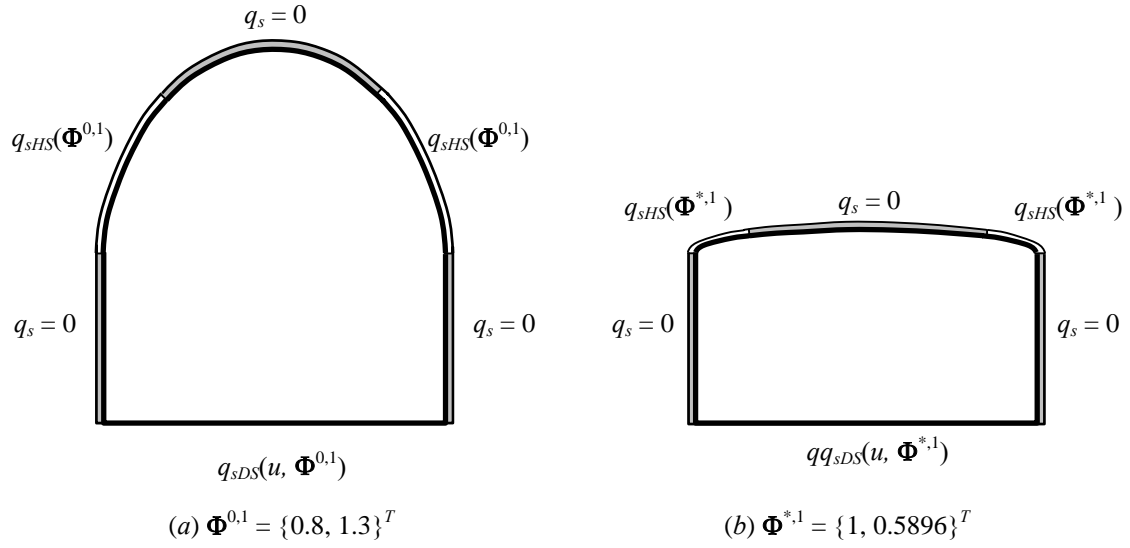


Figure 3.11: Enclosure geometries corresponding to (a)  $\Phi^{0,1} = \{0.8, 1.3\}^T$  and

(b)  $\Phi^{*,1} = \{1, 0.5896\}^T$ .

From examining Figs. 3.8 and 3.10, it is clear that there exists at least one other local minimum at the lower left-hand corner of the feasible region. A second Newton minimization is performed starting from  $\Phi^{0,2} = \{0.5, 1\}^T$ . Three steps are required to reach a second local minimum  $\Phi^{*,2} = \{0, 0.5\}^T$ , which corresponds to  $F(\Phi^{*,2}) = 2.275 \times 10^{-4}$ . This minimization path is shown in Fig. 3.12, and the initial and final enclosure configuration and heat flux distribution over the design surface are shown in Figs. 3.13 and 3.14.



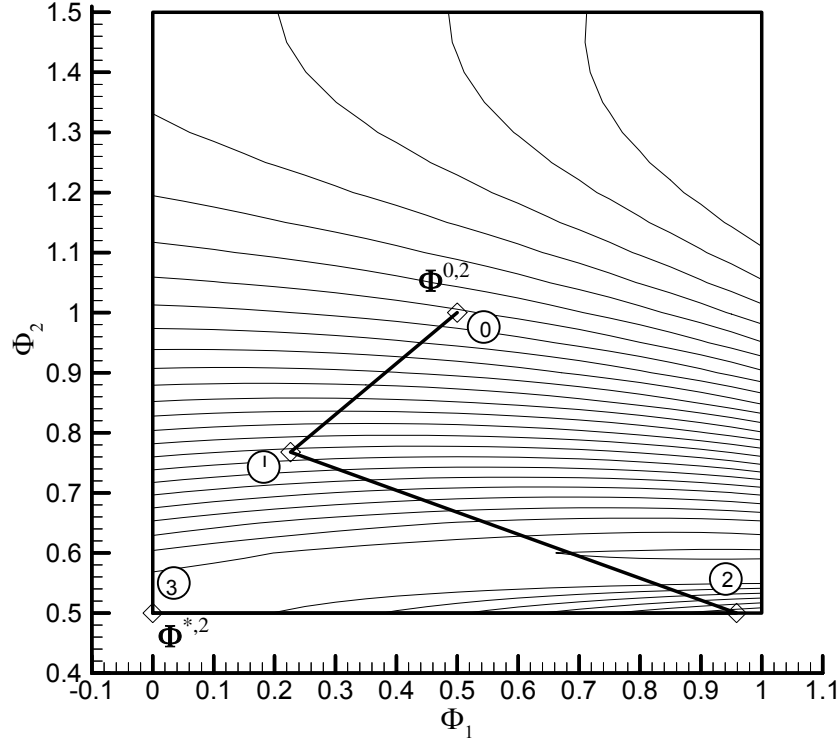


Figure 3.12: Minimization path starting from  $\Phi^{0,2} = \{0.8, 1.3\}^T$  and ending at  $\Phi^{*,2} = \{0, 0.5\}^T$ .

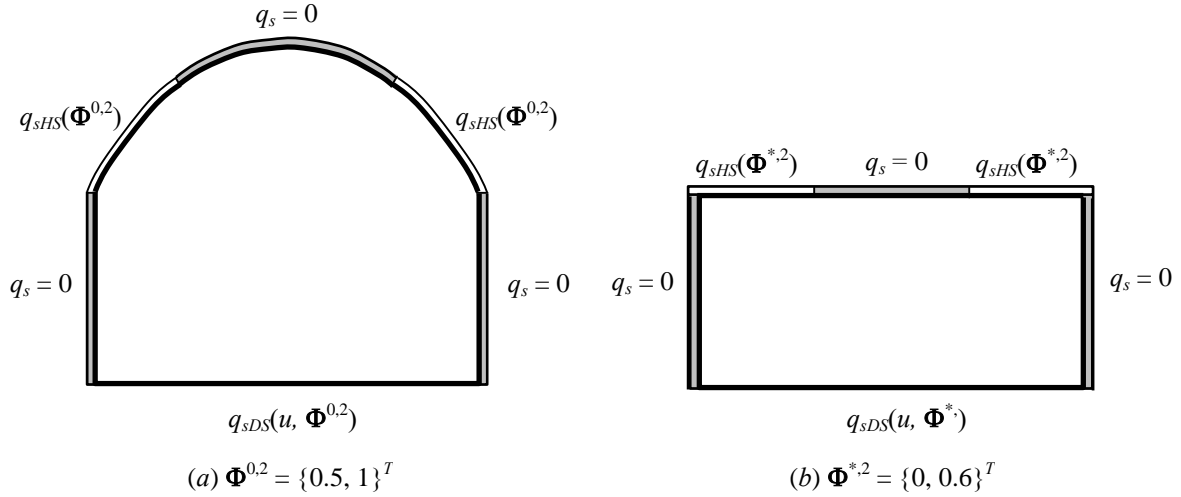


Figure 3.13: Enclosure geometries corresponding to (a)  $\Phi^{0,2} = \{0.5, 1\}^T$  and (b)  $\Phi^{*,2} = \{0, 0.5\}^T$ .

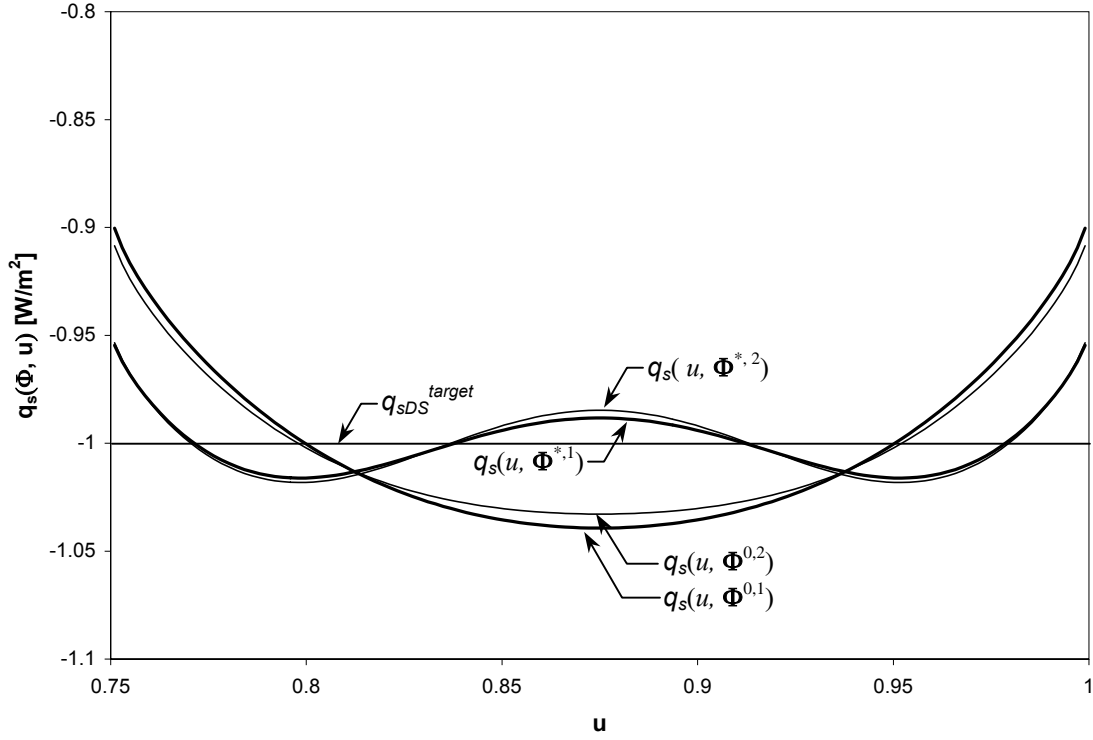


Figure 3.14: Heat flux distributions corresponding to the initial and optimal enclosure configurations shown in Fig. 3.11 and Fig. 3.13.

From examining Fig. 3.14, the heat flux distributions corresponding to the local minima,  $q_s(u, \Phi^{*,1})$  and  $q_s(u, \Phi^{*,2})$ , both match the desired heat flux distribution over the design surface much better than the initial solutions,  $q_s(u, \Phi^{0,1})$  and  $q_s(u, \Phi^{0,2})$ , respectively. The heat flux distribution over the design surface obtained with  $\Phi^{*,1}$  is slightly better than the one obtained with  $\Phi^{*,2}$ , although the difference between these two heat flux distributions is very small. On the other hand, enclosure geometry corresponding with  $\Phi^{*,2}$  is rectangular and therefore would be much easier to implement than the geometry corresponding with  $\Phi^{*,1}$ , where the top of the enclosure is curved.

This example demonstrates that, rather than relying on the solution obtained from a single local minimum, a superior solution to the design problem can sometimes be found by comparing the merits of solutions obtained at multiple local minima.

It should be noted, however, that while the existence and general location of the local minima was easily determined in this case by examining the plot of the objective function over the feasible region, in problems where the objective function is computationally expensive to calculate the CPU time needed to plot the feasible region may be prohibitive and certainly negates the computational efficiency of the gradient-based minimization. Furthermore, most practical design problems involve more than two parameters, which makes visualization of  $F(\Phi)$  very difficult.

Accordingly, local minima are usually located by performing multiple gradient-based minimizations, each from a different initial point. This process can be quite time-consuming if the feasible region is large; in fact, the probability of locating all the local minima approaches unity only as the number of minimizations becomes infinite. This problem can be overcome by employing a sophisticated “multistart” algorithm (e.g. Ugray et al., 2002), which reduces the number of gradient-based minimizations that need to be performed by employing a sophisticated algorithm to determine the most effective locations of the initial points in the feasible region.

### 3.7 Conclusions

This chapter presented an optimization methodology for solving radiant enclosure design problems where both the enclosure geometry and heat flux distributions over the heater surface are unknown. The heat flux distribution is solved using the infinitesimal-area analysis method, and the minimization is carried out using Newton's method. The first- and second-order objective function sensitivities are found from the heat flux sensitivities, which in turn are efficiently calculated by post-processing the radiosity distribution. A method for representing the enclosure geometry using non-uniform rational B-spline (NURBS) curves was presented; this technique is well suited for radiant enclosure analysis because the geometric terms used to solve the heat flux distribution are readily available, and also presents several important advantages in the context of geometric optimization.

The optimization design methodology was demonstrated by solving a 2-D radiant enclosure composed of diffuse-gray walls and containing a transparent medium. The enclosure walls were represented by Bézier curves, which are a type of NURBS curve. In this example, the feasible region contained two local minima corresponding to two different solutions to the design problem. Although the existence and locations of the local minima were found by plotting the objective function over the feasible region in this example, methods for identifying multiple local minima in more complex problems were briefly discussed.

While the optimization methodology presented here can be used to solve geometric optimization problems involving diffuse-walled radiant enclosures, the next

chapter presents a design methodology for optimizing the shape of radiant enclosures containing specularly-reflecting surfaces, which is a more commonly encountered problem in industrial settings.

**Chapter 4:**  
**Geometric Optimization of Radiant Enclosures**  
**Containing Non-Diffuse Surfaces**

## 4.1 Introduction

The design methodology presented in the previous chapter is well suited for optimizing the geometry and heater settings of radiant enclosures composed of diffusely emitting and reflecting surfaces. In reality, however, the optical properties of most surfaces are directionally-dependent and it is often necessary to take this into account in order to construct a sufficiently accurate heat transfer model. Moreover, many applications use specularly-reflecting (mirror-like) surfaces to achieve the desired heat flux and temperature distributions over the design surface. Examples include solar concentrator-collectors, light boxes in illumination applications, and reflectors used in infrared baking ovens.

Until recently, these problems were almost exclusively solved using the forward design methodology, where an acceptable enclosure design is found by trial-and-error, starting from an initial guess. As previously noted, forward design methodologies typically require many iterations and a substantial amount of design time, and are almost always stopped before the solution has been optimized. This is particularly true in this case, since the complicated nature of radiant exchange between non-diffuse surfaces makes an intuitive understanding of the problem physics particularly elusive. Because of this, the designer rarely knows how to modify the enclosure configuration at each iteration in order to maximize the improvement in design performance. Furthermore, the analysis requires considerably more computational effort to carry out at each design iteration, compared with techniques used to treat diffuse-walled enclosure problems.

Many approximate analytical design techniques have been developed as alternatives to the forward design methodology. Methods based on non-imaging optics analysis have proven particularly popular (Winston, 1991). They were first developed to design apparatuses used in high-energy physics applications, and have subsequently been extended to solve many different types of design problems, including those that arise in illumination (e.g. Gordon, 1993) and solar energy applications (e.g. Chen et al., 2001). The most standard of these techniques is the edge-ray method, which uses a complex mathematical procedure based on analytical geometry and calculus to determine the optimal shape of reflector surfaces, and is most often applied to design concentrators that have the highest possible radiant heat flux concentration ratios between entrance and exit apertures. Güven (1994) also presented a semi-analytical method for designing collector geometries. In this method, the optimal collector geometry is found by first deriving an analytical expression for the intercept factor (defined as the fraction of reflected radiation that reaches the receiver) as a function of two geometric parameters. The collector geometry is then optimized by setting the derivatives of this function with respect to the geometric parameters equal to zero, and then solving the resulting equations. While approximate analytical techniques are very powerful and widely-used design tools, however, they can only be applied to treat a few types of enclosure geometries and do not account for surface properties having both diffuse and specular components, as shown by Maruyama (1993). Moreover, the designer must possess specialized mathematical knowledge in order to carry out the analysis.



Because of these shortcomings, many design methodologies have been developed that are based on numerical simulations. Most of these simulations are based on the Monte Carlo ray-tracing method; this powerful analysis technique can be applied to solve complex problems and is also very straightforward to implement because of its inherent simplicity. For example, Ryan et al. (1998) used a Monte Carlo technique to analyze a cylindrical solar collector and drew general conclusions about the optimal collector configuration by performing a series of univariate parametric studies. Mushaweck et al. (2000) calculated the optimal reflector shapes for non-tracking parabolic trough concentrators. The average utilizable power was first calculated by performing a numerical analysis involving ray-tracing for concentrators with different upper and lower acceptance angles, and then plotted over a rectangular feasible region defined by the maximum and minimum values of the acceptance angles in order to identify the optimal reflector shape. Although numerical simulation techniques can treat a more extensive set of problems than the approximate analytical methods, both of the above studies relied on primitive optimization algorithms that required a substantial amount of design time and restricted the number of design parameters that could be considered in the analysis.

A more sophisticated optimization methodology is described by Ashdown (1994), in which a ray-tracing technique used to simulate illumination within an enclosure is coupled with a genetic algorithm that searches for the globally optimal enclosure geometry. Genetic algorithms mimic natural selection as it occurs in nature. This class of algorithms generates new designs by “mating” pairs of previously generated designs and by “mutating” existing designs. The designs that perform well are favored in the

minating process, and a near-optimum solution is usually found after many generations. (Forrest, 1993, provides a detailed discussion of genetic algorithms.)

This chapter describes an optimization design methodology originally presented in Daun et al. (2003d) that overcomes many drawbacks of the standard numerical and approximate analytical approaches that are traditionally applied to design enclosures containing non-diffuse surfaces. The objective of the design methodology is to identify the enclosure geometry and heater settings that produce the desired heat flux and temperature distribution over the design surfaces. As is done in the other optimization methodologies, the desired temperature distribution is first specified over the design surface and the enclosure configuration that produces the desired heat flux distribution is then found by minimizing an objective function. A Monte Carlo technique based on exchange factors is used to calculate the heat flux distribution over the design surface; unlike the infinitesimal-area analysis used in the previous chapters, this technique induces a statistical uncertainty in the objective function and gradient vector. Accordingly, the Kiefer-Wolfowitz method, a gradient-based algorithm developed for optimizing stochastic systems in which unbiased gradient estimates are unavailable, is employed to carry out the minimization. The optimization methodology is demonstrated by using it to solve two design problems involving two-dimensional radiant enclosures containing both diffuse and non-diffuse surfaces.

## 4.2 Optimization Strategy

Consider the radiant enclosure design problem shown in Fig. 4.1, where some of the enclosure surfaces have directionally dependent properties. The objective of this problem is to determine the optimum enclosure geometry and heat flux distribution over the heater surface that produces the desired temperature and heat flux distribution over the design surface. As was done in the previous chapter, the design problem is solved by specifying the temperature distribution over the design surface and using the unspecified heat flux distribution to define the objective function,

$$F(\Phi) = \frac{1}{N_{DS}} \sum_{j=1}^{N_{DS}} [q_{sj}(\Phi) - q_{sj}^{target}]^2, \quad (4.1)$$

where the design parameters contained in  $\Phi$  control the enclosure geometry and the heat flux distribution over the heater surface. The enclosure configuration that produces a heat flux distribution most closely matching the desired distribution corresponds to the design parameters contained in  $\Phi^*$ , which in turn is found by minimizing  $F(\Phi)$ .

Unlike the methodology presented in the previous chapter, however, the complicated nature of radiant exchange between non-diffuse surfaces precludes the use of a deterministic technique like infinitesimal-area analysis to solve for the heat flux distribution at each iteration. Instead, the heat flux distribution over the design surface is estimated using exchange factors calculated by the Monte Carlo method. As will be shown in the following section, the heat flux at each discrete location over the design surface,  $q_{sj}(\Phi)$ , is estimated by  $\tilde{q}_{sj}(\Phi)$ , which contains a statistical uncertainty due to sampling errors in the exchange factors. Accordingly, the objective function in Eq. (4.1) is approximated by

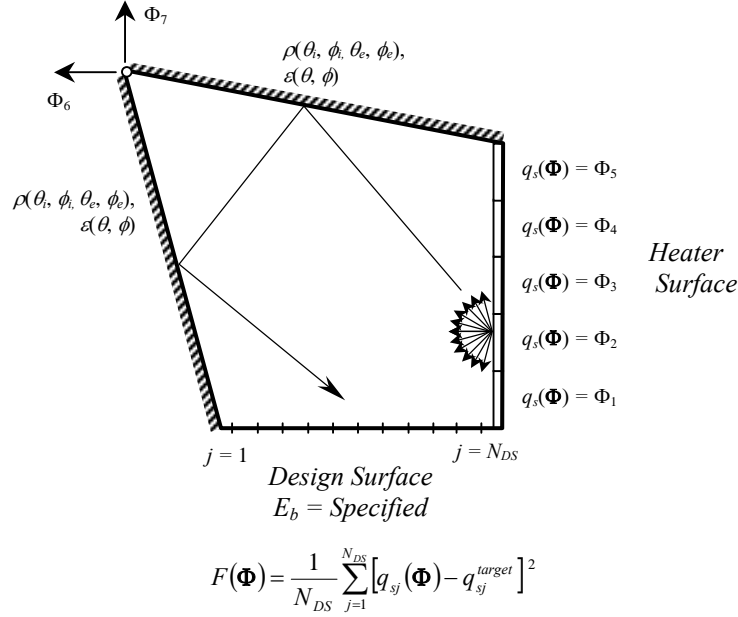


Figure 4.1: Example radiant enclosure design problem, where some surfaces have directionally-dependent properties.

$$F(\Phi) \approx \tilde{F}(\Phi) = \frac{1}{N_{DS}} \sum_{j=1}^{N_{DS}} [\tilde{q}_{sj}(\Phi) - q_{sj}^{target}]^2, \quad (4.2)$$

which is subject to a sampling error,  $\delta_1(\Phi)$ , induced by the statistical uncertainty in the heat flux distribution. Using the technique described by Hammersley and Handscomb (1992), the sampling error is estimated by

$$\delta_1(\Phi) \approx \left[ \sum_{i=1}^{N_{DS}} \sum_{j=1}^{N_{DS}} \frac{\partial \tilde{F}(\Phi)}{\partial \tilde{q}_{si}} \frac{\partial \tilde{F}(\Phi)}{\partial \tilde{q}_{sj}} \Gamma_{ij}(\Phi) \right]^{1/2}, \quad (4.3)$$

where  $\Gamma(\Phi)$  is the variance-covariance matrix of the heat flux values and is defined in the next section. The uncertainty inherent in the evaluation of  $\tilde{F}(\Phi)$  makes the optimization of stochastic systems somewhat more complicated than that of deterministic systems, since the “exact” value of  $F(\Phi)$  is unknown.

Many methods used to optimize stochastic systems are based on those used to optimize deterministic systems. The Kiefer-Wolfowitz method (Kiefer and Wolfowitz, 1952, Kushner and Clark, 1978) is a gradient-based optimization technique used when an unbiased estimator of  $\nabla F(\Phi)$  is unavailable. This method is based on the steepest-descent algorithm described in Appendix A; at the  $k^{th}$  iteration, the step size is set equal to a non-vanishing sequence based on the iteration number,

$$\alpha^k = \frac{\alpha_0}{k^a}, \quad 0 \leq a \leq 1, \quad (4.4)$$

and the search direction is found by

$$\mathbf{p}^k = -\frac{\tilde{\mathbf{g}}(\Phi^k)}{\|\tilde{\mathbf{g}}(\Phi^k)\|}, \quad (4.5)$$

where  $\tilde{\mathbf{g}}(\Phi^k)$  is a second-order central difference approximation of  $\nabla F(\Phi^k)$ . The  $p^{th}$  term of  $\tilde{\mathbf{g}}(\Phi^k)$  is given by

$$\tilde{g}_p(\Phi^k) = \frac{\tilde{F}(\Phi^k + \mathbf{e}_p \cdot h^k) - \tilde{F}(\Phi^k - \mathbf{e}_p \cdot h^k)}{2h^k}, \quad (4.6)$$

where  $\mathbf{e}_p$  is the unit vector of the  $p^{th}$  direction, and  $h^k$  is the interval used in the finite difference approximation at the  $k^{th}$  iteration.

This estimate contains two sources of error: a bias error,  $\delta_{2,p,k}(\Phi^k, h^k)$ , due to the truncation of the higher-order terms in the finite difference approximation, and a random error,  $\delta_{3,p,k}(\Phi^k, h^k)$ , induced by the sampling error in  $\tilde{F}(\Phi)$ . (This sampling error tends to dominate the finite difference approximation of the higher order derivatives, which is why the steepest-descent direction is used instead of the Newton or quasi-Newton directions.) The bias error is given by (Pflug, 1996)

$$\delta_{2,p,k}(\Phi^k, h^k) = \nabla F_p(\Phi^k) - \frac{\tilde{F}(\Phi^k + \mathbf{e}_p \cdot h^k) - \tilde{F}(\Phi^k - \mathbf{e}_p \cdot h^k)}{2h^k}. \quad (4.7)$$

(Note that in this application the exact value of the gradient vector is unknown.) Since the central difference approximation improves with diminishing step size,  $\delta_{2,p,k}(\Phi^k, h^k)$  decreases as  $h^k$  becomes small. Assuming independent estimates of  $\tilde{F}(\Phi^k + \mathbf{e}_p \cdot h^k)$  and  $\tilde{F}(\Phi^k - \mathbf{e}_p \cdot h^k)$ , the random error is found from

$$\delta_{3,p,k}(\Phi^k, h^k) = \frac{[\delta_1^2(\Phi^k + \mathbf{e}_p \cdot h^k) + \delta_1^2(\Phi^k - \mathbf{e}_p \cdot h^k)]^{1/2}}{h^k}, \quad (4.8)$$

and tends to *increase* as  $h^k$  becomes small, since decreasing  $h^k$  does not necessarily decrease the magnitude of the numerator in Eq. (4.8). It is therefore important to select an intermediate value for  $h^k$  that ensures that both  $\delta_{2,p,k}(\Phi^k, h^k)$  and  $\delta_{3,p,k}(\Phi^k, h^k)$  are sufficiently small; this is particularly true in the vicinity of  $\Phi^*$ , where the magnitude of the gradient vector becomes small and these errors begin to dominate the estimate of  $\nabla F(\Phi^k)$ . One choice is to reduce  $h^k$  with each successive iteration according to a series similar to Eq. (5.4),

$$h^k = \frac{h^0}{k^b}, \quad 0 \leq b \leq 1. \quad (4.9)$$

Pflug (1996) recommends values of  $a = 1$  and  $b = 1/3$  for Eqs. (4.4) and (4.9), respectively.

Another way of reducing  $\delta_{3,p,k}(\Phi^k, h^k)$  is to use the same sequence of random numbers (common random numbers) to generate both  $\tilde{F}(\Phi^k + \mathbf{e}_p \cdot h_k)$  and  $\tilde{F}(\Phi^k - \mathbf{e}_p \cdot h_k)$ ; this approach is discussed in greater detail in Morton and Popova (2001). Finally, the random error can be reduced as  $\Phi^k$  approaches  $\Phi^*$  by increasing number of statistical samples used to approximate the objective function at each

successive iteration, which decreases both  $\delta_1(\Phi^k + \mathbf{e}_p \cdot h_k)$  and  $\delta_1(\Phi^k - \mathbf{e}_p \cdot h_k)$ . This technique is applied to solve the example problems described later in this chapter.

### 4.3 Solution of the Governing Equations

In this application, the heat flux distribution over the design surface is estimated using exchange factors calculated by the Monte Carlo method. The Monte Carlo method has been extensively used to analyze many different radiant enclosure problems, and is especially well suited for analyzing radiant enclosures containing surfaces with directionally-dependent properties. Although the Monte Carlo method has been described extensively in the literature (e.g. Siegel and Howell, 2002a, Modest, 2003) it is presented again here in order to demonstrate its implementation in the context of stochastic optimization.

The first step of this method is to discretize the enclosure surface into  $N$  elements, with the  $i^{th}$  element having an area  $\Delta A_i$ . The emissivity,  $\varepsilon_i$ , and either the heat flux or the temperature are known at every point on the enclosure surface and are taken to be uniform over each element. In particular, it is assumed that the desired temperature distribution is specified for elements located on the design surface,  $T_i = T_i^{target}$ , and the heat flux distribution corresponding with the heater settings contained in  $\Phi$ ,  $q_{si}(\Phi)$ , is specified for elements on the heater surface. Performing an energy balance over the  $i^{th}$  element results in the equation

$$q_{si}(\Phi) + \sum_{j=1}^N \varepsilon_j E_{bj}(\Phi) \mathcal{F}_{ji}(\Phi) \frac{\Delta A_j(\Phi)}{\Delta A_i(\Phi)} = \varepsilon_i E_{bi}(\Phi), \quad (4.10)$$

where  $E_{bi}(\Phi) = \sigma T_i^4(\Phi)$  and  $\mathcal{F}_{ji}(\Phi)$  is the exchange factor from the  $j^{th}$  to the  $i^{th}$  element, defined as the fraction of the total radiant energy emitted by the  $j^{th}$  element that is absorbed by the  $i^{th}$  element. By applying the reciprocity rule of exchange factors, Eq. (4.10) can be rewritten in a more compact form,



$$\frac{q_{si}(\Phi)}{\varepsilon_i} + \sum_{j=1}^N E_{bj}(\Phi) \mathcal{F}_{ji}(\Phi) = E_{bi}(\Phi). \quad (4.11)$$

Although the analytical solution for the exchange factor is tractable for diffuse-walled enclosure problems, this is not the case if the enclosure surface properties are directionally-dependent. Instead, the exchange factor,  $\mathcal{F}_{ij}(\Phi)$ , can be expressed as the expectation of a random variable,  $E[\mathcal{F}_{ij}(\Phi, \xi)]$ , where  $\xi$  contains three other random variables that specify the emission location and direction of a random bundle leaving the  $i^{th}$  surface element, as shown in Fig. 4.2. Theoretically,  $\mathcal{F}_{ij}(\Phi)$  could be found by integrating over the probability distributions governing each  $\xi_i$ ,

$$\mathcal{F}_{ij}(\Phi) = E[\mathcal{F}_{ij}(\Phi, \xi)] = \int_0^1 \int_0^1 \int_0^1 \mathcal{F}_{ij}(\Phi, \xi) dP_1(\xi_1) dP_2(\xi_2) dP_3(\xi_3). \quad (4.12)$$

(This is, in fact, equivalent to carrying out the integration used to calculate exchange factors between diffuse surfaces.) Instead, we use a Monte Carlo simulation to estimate  $\mathcal{F}_{ij}(\Phi)$  in Eq. (4.12),

$$E[\mathcal{F}_{ij}(\Phi, \xi)] \approx \tilde{\mathcal{F}}_{ij}(\Phi) = \frac{1}{N_{bi}} \sum_{m=1}^{N_{bi}} \mathcal{F}_{ij}(\Phi, \xi^m) = \frac{N_{bij}}{N_{bi}}, \quad (4.13)$$

where  $N_{bi}$  is the total number of bundles emitted from the  $i^{th}$  element, and  $N_{bij}$  is the number of those bundles absorbed by the  $j^{th}$  element. Due to the law of large numbers, the Monte Carlo approximation of  $E[\mathcal{F}_{ij}(\Phi, \xi)]$  becomes exact with probability one as  $N_{bi}$  approaches infinity. Since we are restricted to using a finite number of bundles, however,  $\tilde{\mathcal{F}}_{ij}(\Phi)$  contains a random error that propagates throughout the solution.

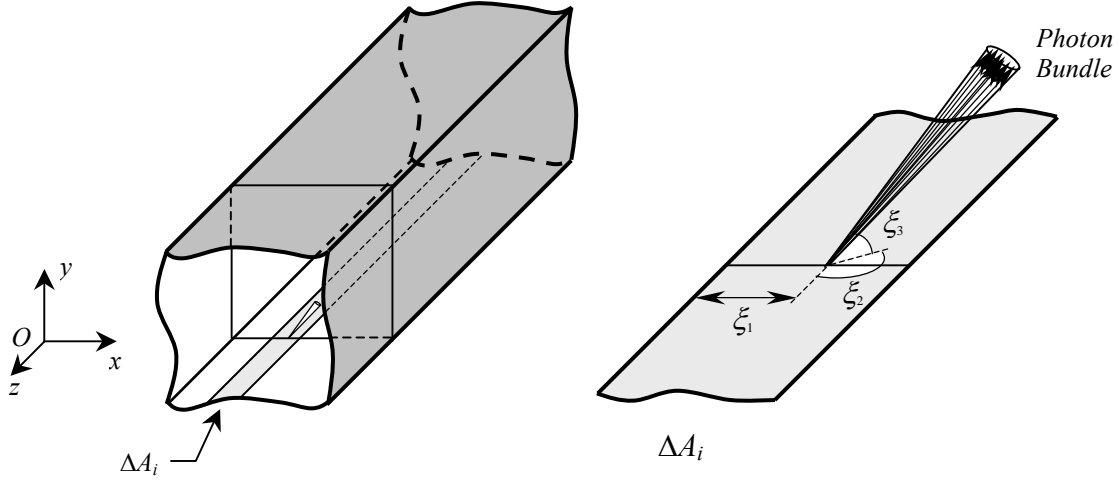


Figure 4.2: Random variables contained in  $\xi$  control the location and direction of bundle emission from a surface element in a two-dimensional enclosure.

Assume that the elements are reordered so that  $T_i$  is specified for  $i = 1 \dots N'$  and  $q_{si}$  is specified for  $i = N'+1 \dots N$ . Equation (4.11) can then be rewritten as

$$\frac{\tilde{q}_{si}(\Phi)}{\epsilon_i} + \sum_{j=N'+1}^N \tilde{E}_{bj}(\Phi) \tilde{\mathcal{F}}_{ij}(\Phi) = E_{bi}(\Phi) - \sum_{j=1}^{N'} E_{bj}(\Phi) \tilde{\mathcal{F}}_{ij}(\Phi), \quad (4.14)$$

for elements where  $T_i$  is specified, and

$$\tilde{E}_{bi}(\Phi) - \sum_{j=N'+1}^N \tilde{E}_{bj}(\Phi) \tilde{\mathcal{F}}_{ij}(\Phi) = \frac{q_{si}(\Phi)}{\epsilon_i} + \sum_{j=1}^{N'} E_{bj}(\Phi) \tilde{\mathcal{F}}_{ij}(\Phi), \quad (4.15)$$

for elements where  $q_{si}$  is specified. Equations (4.14) and (4.15) are arranged so that the  $T_i$  and  $q_{si}$  terms on the right-hand side are known, while those on the left are unknown and will contain a random error when they are solved due to the uncertainty in the exchange factors. Writing these equations for all elements results in a system of  $N$  equations containing  $N$  unknowns that can be rearranged into a matrix equation,

$$\tilde{\mathbf{A}}(\Phi) \tilde{\mathbf{x}}(\Phi) = \tilde{\mathbf{b}}(\Phi), \quad (4.16)$$

where  $\tilde{\mathbf{x}}(\Phi) = \{\tilde{q}_{s1}(\Phi), \dots, \tilde{q}_{sN'}(\Phi), \tilde{E}_{bN'+1}(\Phi), \dots, \tilde{E}_{bN}(\Phi)\}^T$ . (The rank of  $\tilde{\mathbf{A}}(\Phi)$  can be reduced by excluding equations corresponding to elements where  $q_{si}$  is equal to zero.) This linear system is usually well conditioned, and Eq. (4.16) can be easily solved to yield the heat flux distribution over the design surface.

The heat flux distribution over the design surface is subject to a random error induced by the sampling error in the exchange factors (which is responsible for  $\delta_i(\Phi)$ ), as well as a bias error caused by assuming a uniform heat flux and temperature distribution over each element. The former error is reduced by increasing the number of bundles used to estimate the exchange factors, while the latter diminishes at higher levels of grid refinement. Nevertheless, both errors result in a grid-dependent objective function containing a statistical uncertainty, which renders it difficult to optimize.

The magnitude of the random error is estimated by performing a replication procedure. Suppose a total of  $N_b$  bundles are used to calculate the exchange factors throughout the process. A sequence of Monte Carlo simulations are used to estimate  $p$  independent sets of exchange factors, each using  $N_b/p$  bundles. Each set of exchange factors is then used to find an estimate of the heat flux distribution over the design surface, through Eq. (4.16). Performing this procedure for each set of exchange factors results in  $p$  independent solutions,  $\tilde{\mathbf{x}}_m$ ,  $m = 1 \dots p$ .

The heat flux at each discrete point over the design surface is then approximated as the average of the  $p$  independent solutions,

$$\tilde{q}_{si}(\Phi) = \frac{1}{p} \sum_{m=1}^p \tilde{q}_{si}^m(\Phi), \quad (4.17)$$

while the corresponding random error associated with  $q_s(\Phi)$  is estimated from the sample standard deviation,

$$\sigma_{i,ave}(\Phi) = \frac{\sigma_i(\Phi)}{\sqrt{p}}, \quad (4.18)$$

where  $\sigma_i^2(\Phi)$  is the sample variance of the  $p$  measurements, given by

$$\sigma_i^2(\Phi) = \frac{1}{p-1} \sum_{m=1}^p [\tilde{q}_{si}^m(\Phi) - \tilde{q}_{si}(\Phi)]^2. \quad (4.19)$$

It is important to note that while the random error of each exchange factor is independent, the operations performed when solving Eq. (4.16) induces a weak interdependence between the random errors of the heat flux terms. The degree of dependence between the random errors in  $\tilde{q}_{si}(\Phi)$  and  $\tilde{q}_{sj}(\Phi)$  is estimated from their covariance,

$$\sigma_{ij}^2(\Phi) = \frac{1}{p-1} \sum_{m=1}^p [\tilde{q}_{si}^m(\Phi) - \tilde{q}_{si}(\Phi)][\tilde{q}_{sj}^m(\Phi) - \tilde{q}_{sj}(\Phi)]. \quad (4.20)$$

The sample variance and covariance terms form the variance-covariance matrix,  $\Gamma(\Phi)$ , which in turn is used to estimate the random error in  $\tilde{F}(\Phi)$  through Eq. (4.3).

#### 4.4 Implementation and Demonstration

The optimization methodology presented in the previous sections is demonstrated by using it to design two radiant enclosures. In both cases, the goal of the design process is to identify an enclosure geometry that produces a uniform heat flux distribution over the temperature-specified design surface.

The first problem involves an enclosure consisting of a heater surface, a temperature-specified design surface, and two adiabatic reflector surfaces, and is shown in Fig. 4.3. The heater and design surfaces are both diffuse and have an emissivity of  $\varepsilon = 1$ , while the adiabatic surfaces are perfectly specular with a reflectivity  $\rho_s = 1$ . A uniform heat input  $q_{SHS} = 1 \text{ W/m}^2$  is maintained over the heater surface, and the design surface has an emissive power  $E_{bDS} = 0 \text{ W/m}^2$ .

The objective of the design problem is to find the enclosure geometry that produces a uniform heat flux of  $q_s^{target} = -1 \text{ W/m}^2$  over the design surface, which is done by minimizing the objective function defined in Eq. (4.2). The design parameters contained in  $\Phi = \{\Phi_1, \Phi_2\}^T$  control the  $x$ - and  $y$ -coordinates of the upper left-hand vertex of the enclosure, respectively. Minimization is carried out using the Kiefer-Wolfowitz method starting from an arbitrarily chosen initial point,  $\Phi^0 = \{-0.5, 0.5\}^T$ .

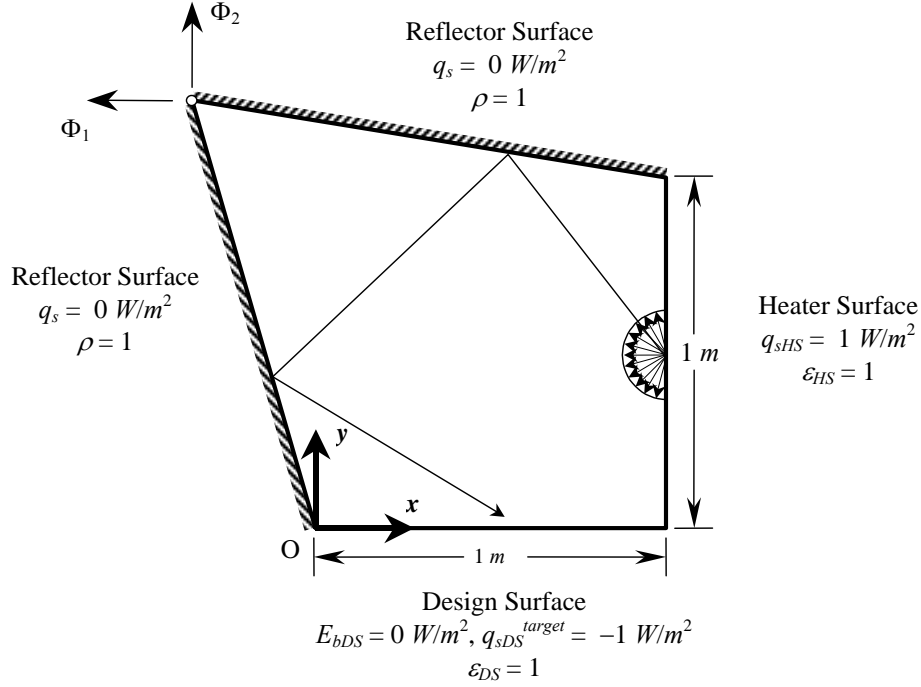


Figure 4.3: Radiant enclosure design problem involving two adiabatic, specularly-reflecting surfaces.

As previously mentioned, assuming a uniform heat flux and emissive power distribution over each surface element produces a discretization error in  $\tilde{q}_{si}(\Phi)$ , and accordingly in  $\tilde{F}(\Phi)$ , that diminishes with increasing levels of grid refinement. Accordingly, a grid refinement study was performed at  $\Phi^0 = \{-0.5, 0.5\}^T$  in order to determine how many elements should be used to carry out the analysis, which is shown in Fig. 4.4. A constant ratio of bundles to elements,  $N_b/N = 2 \times 10^5$ , was used to obtain approximately the same sampling error at every level of grid refinement. An estimate of the grid-independent solution is obtained from the highest level of grid refinement,

$\tilde{F}_\infty(\Phi^0) = \tilde{F}_{N=2048}(\Phi^0) = 0.1690$ . The discretization error at the next highest level of refinement,  $N = 1024$ , is then estimated as  $-1.6\%$  of the grid-independent solution.

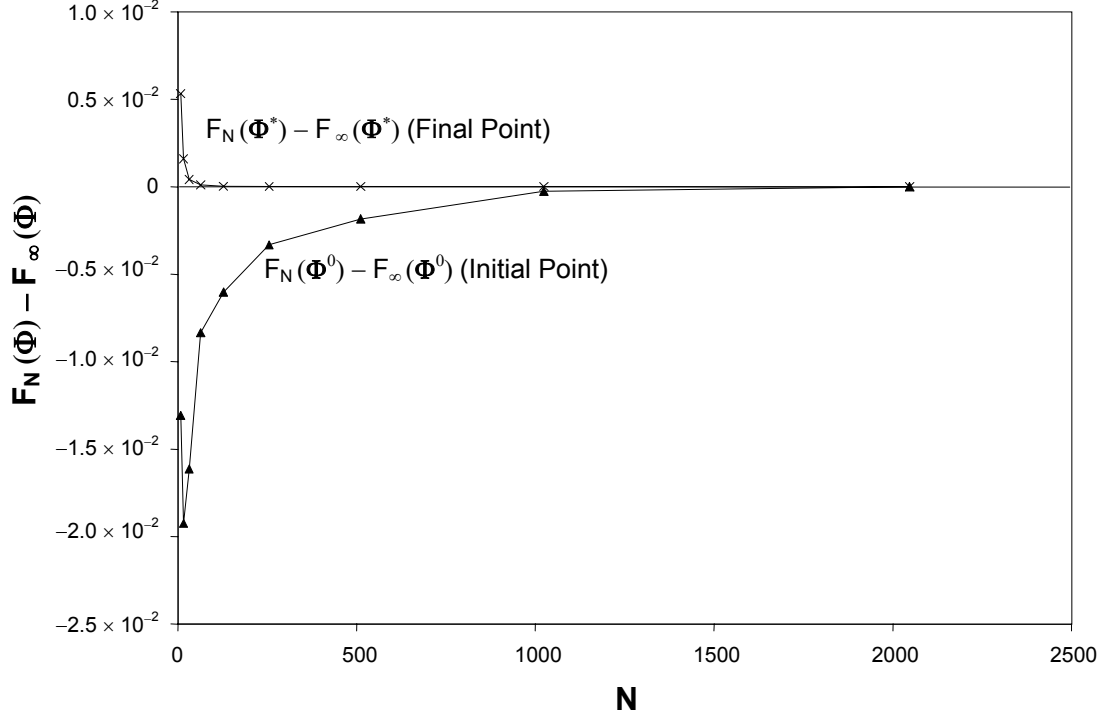


Figure 4.4: Grid refinement study for the design problem of Fig. 4.3.

Next, the effect of the number of bundles on the random error,  $\delta_1(\Phi)$ , is demonstrated by systematically increasing the number of bundles used to calculate  $\tilde{F}(\Phi^0)$  for an enclosure having  $N = 1024$  surface elements. As shown in Fig. 4.5, a power law relationship exists between the number of bundles and the random error,

$$\delta_1(\Phi^0) = 1.4320 N_b^{-0.5272}, \quad (4.21)$$

which is consistent with the  $1/\sqrt{N_b}$  trend predicted by the central limit theorem.

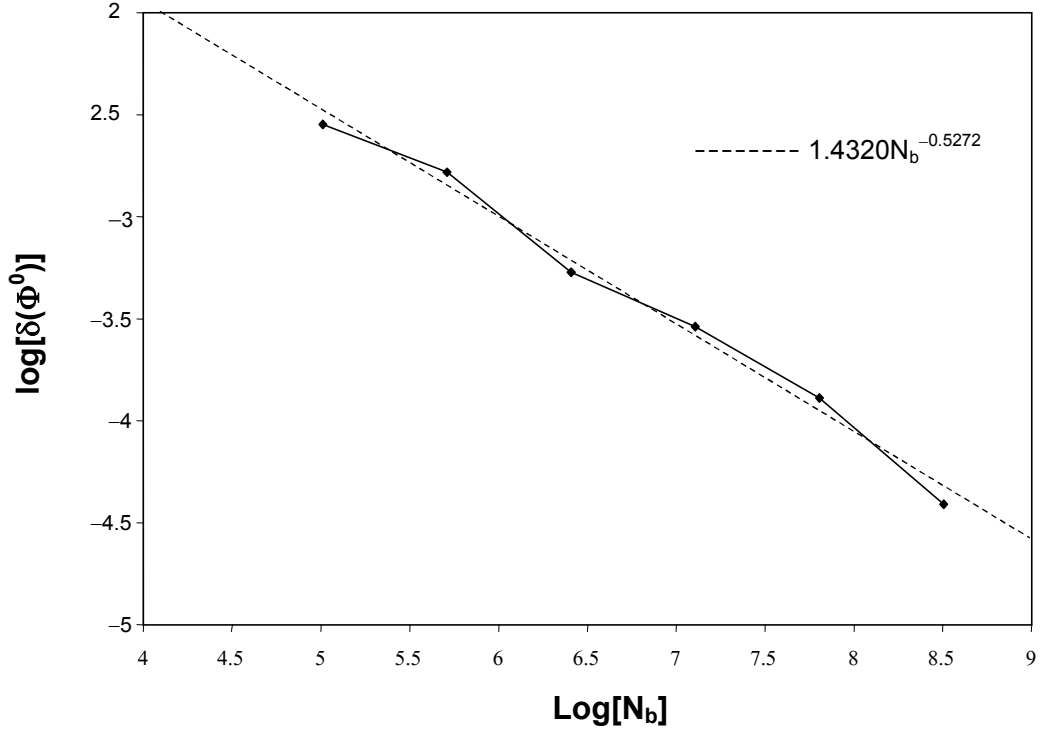


Figure 4.5: Effect of the number of bundles on  $\delta_1(\Phi^0)$  for the design problem of Fig. 4.3.

Based on the above results, the Kiefer-Wolfowitz minimization was carried out using  $N = 1024$  surface elements. As previously noted, as  $\Phi^k$  approaches  $\Phi^*$  the magnitude of the gradient vector becomes smaller and the search direction is more susceptible to influence from  $\delta_1(\Phi^k, h^k)$ . To account for this effect, the number of bundles used in the simulation was increased with each successive iteration according to

$$N_b = A \log_{10} k + B, \quad (4.22)$$



where  $A = 160 \times 10^6$  and  $B = 20 \times 10^6$ , chosen based on the results of the grid and bundle refinement studies shown in Figs. 4.4 and 4.5. The minimization procedure was stopped when  $|\tilde{F}(\Phi^k) - \tilde{F}(\Phi^{k-1})| \leq 10^{-4}$ .

The resulting solution path is shown in Fig. 4.6. Seven steps were required to identify a local minimum at  $\Phi^* = \{0.0034, 0.8457\}^T$ , with  $\tilde{F}(\Phi^*) = 2.26 \times 10^{-4}$ . The enclosure geometries corresponding to  $\Phi^0$  and  $\Phi^*$  are shown in Fig. 4.7, while the heat flux distributions over the design surface are shown in Fig. 4.8. A grid refinement study on  $\tilde{F}(\Phi^*)$  is also shown in Fig. 4.4, which confirms that a sufficient level of grid refinement was used to find the optimal solution.

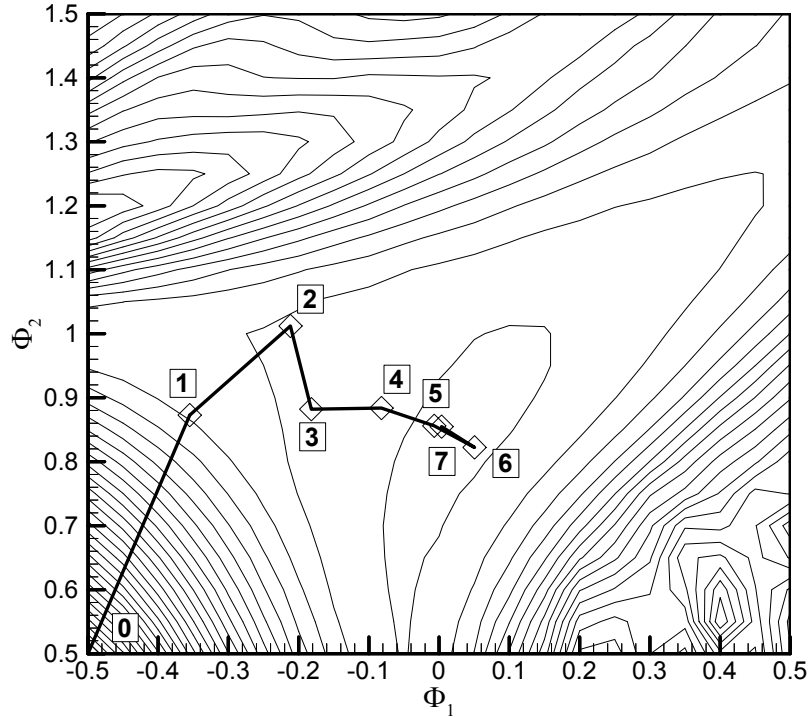


Figure 4.6: Minimization path for the design problem of Fig. 4.3.

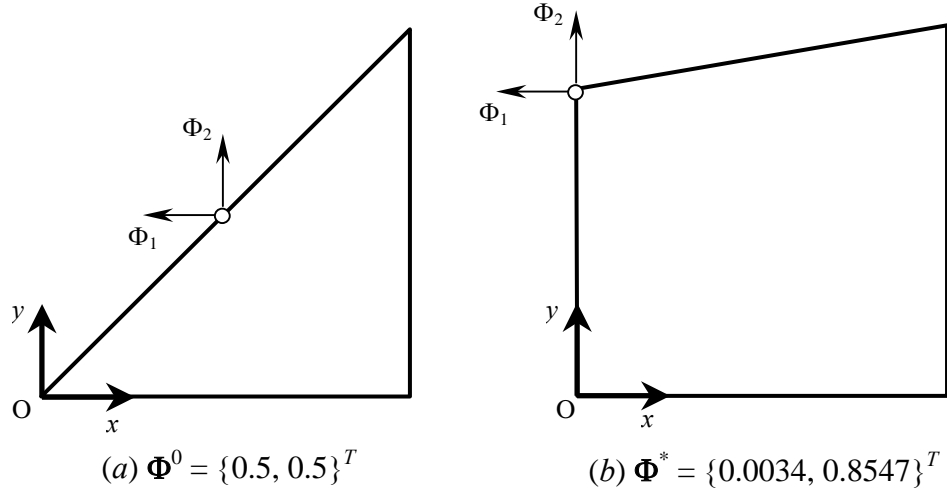


Figure 4.7: Initial and final enclosure geometries for the design problem of Fig. 4.3.

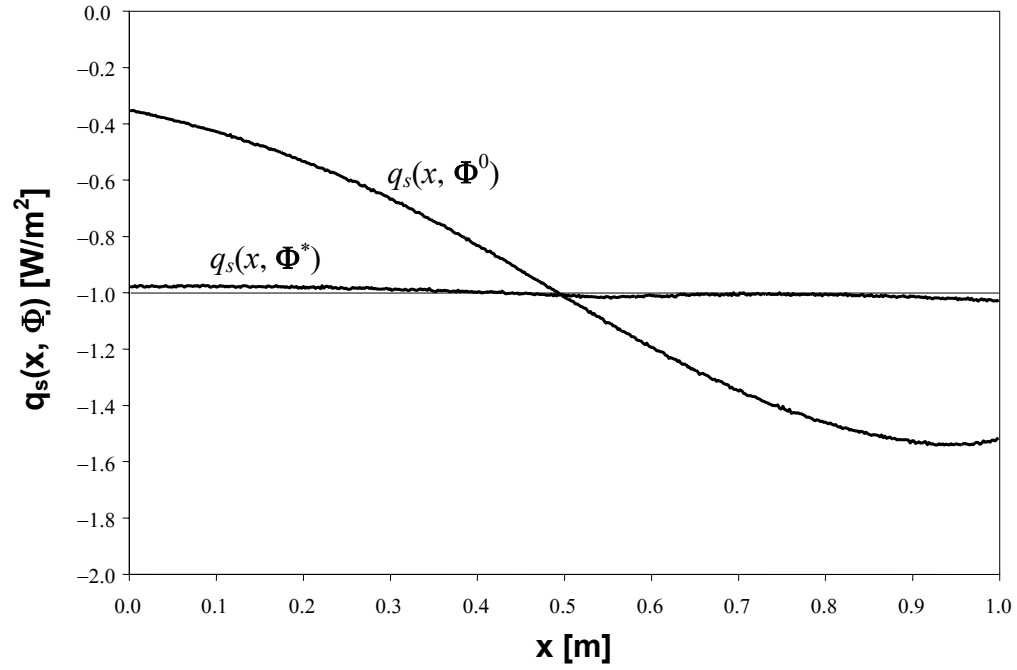


Figure 4.8: Initial and final design surface heat flux distributions for the design problem of Fig. 4.3.

The second design problem is shown in Fig. 4.9, and is similar to the imaging furnace described by Maruyama (1993). The enclosure consists of a cylindrical heating element surrounded by six reflecting surfaces and a design surface. The heater surface is black and diffuse, and has a specified heat flux of  $q_{sHS} = 1/2\pi R \text{ W/m}^2$ , where the radius of the heater element is  $R = 0.1 \text{ m}$ . The design surface has a specified emissive power  $E_b^{target} = 0 \text{ W/m}^2$  and an emissivity  $\varepsilon_{DS} = 0.6$ . The reflecting surfaces have the properties of polished nickel, with  $\rho_s = 0.65$ ,  $\rho_d = 0.25$ , and  $\varepsilon = 0.1$  (Birkebak et al. 1964.)

The objective of this problem is to identify the enclosure geometry that most closely produces a heat flux of  $q_s^{target} = -0.5 \text{ W/m}^2$  over the design surface. The enclosure geometry is governed by five different parameters:  $\Phi_1$  through  $\Phi_4$  control the orientation of the reflector surfaces, while  $\Phi_5$  specifies the height of the heater element over the design surface. The optimal enclosure configuration was determined by minimizing the objective function defined in Eq. (4.2) using the Kiefer-Wolfowitz method.

Based on the results of grid- and bundle refinement studies similar to the ones shown in Figs. 4.4 and 4.5, 1024 surface elements were again used throughout the analysis and the values of  $A$  and  $B$  in Eq. (4.22) were set equal to  $6 \times 10^6$  and  $5 \times 10^5$ , respectively. Minimization was started at two different points,  $\Phi^{0,1} = \{0.25, 0.75, 0.75, 1.25, 0.5\}^T$ , and  $\Phi^{0,2} = \{1, 1, 1, 1, 0.5\}^T$ , and was stopped when  $|\tilde{F}(\Phi^k) - \tilde{F}(\Phi^{k-1})| < 10^{-6}$ . Two different local minima were identified; starting from  $\Phi^{0,1}$ , 18 steps were required to reach the first local minimum at  $\Phi^{*,1} = \{0.2982, 0.6980,$

$0.7600, 1.3955, 0.5588\}^T$  with  $\tilde{F}(\Phi^{*,1}) = 7.87 \times 10^{-5}$ . Starting from  $\Phi^{0,2}$ , 29 steps were required to reach  $\Phi^{*,2} = \{0.9453, 0.8777, 1.03397, 1.0184, 0.5520\}^T$  with  $\tilde{F}(\Phi^{*,2}) = 8.07 \times 10^{-5}$ .

The initial and final enclosure geometries for these cases are shown in Figs. 4.10 and 4.11, and the corresponding heat flux distributions over the design surface is shown in Fig. 4.12. Figure 4.12 shows that the heat flux distribution over the design surface produced by the enclosure geometries corresponding to  $\Phi^{*,1}$  and  $\Phi^{*,2}$  are more uniform than those produced by the initial enclosure geometries even though the initial and optimal enclosure geometries are very similar in both cases, as shown in Figs. 4.10 and 4.11. Because the heat flux distribution over the design surface is sensitive to small perturbations in the enclosure geometry, this design problem would be particularly difficult to solve using the forward “trial-and-error” methodology.

These results also suggest the existence of many local minima over the feasible region, each one corresponding to a unique enclosure geometry and heat flux distribution over the design surface. In order to find the global minimum, or at least one of the smaller local minima, it may be necessary to use a technique better suited to minimizing multimodal objective functions, such as a heuristic multistart algorithm (e.g. Ugray et al., 2002) or a metaheuristic optimization algorithm. Because these types of techniques work by evaluating the objective function at many different locations over the feasible region, however, the computational effort required to carry out these procedures to design this class of radiant enclosure would be considerable and potentially prohibitive.

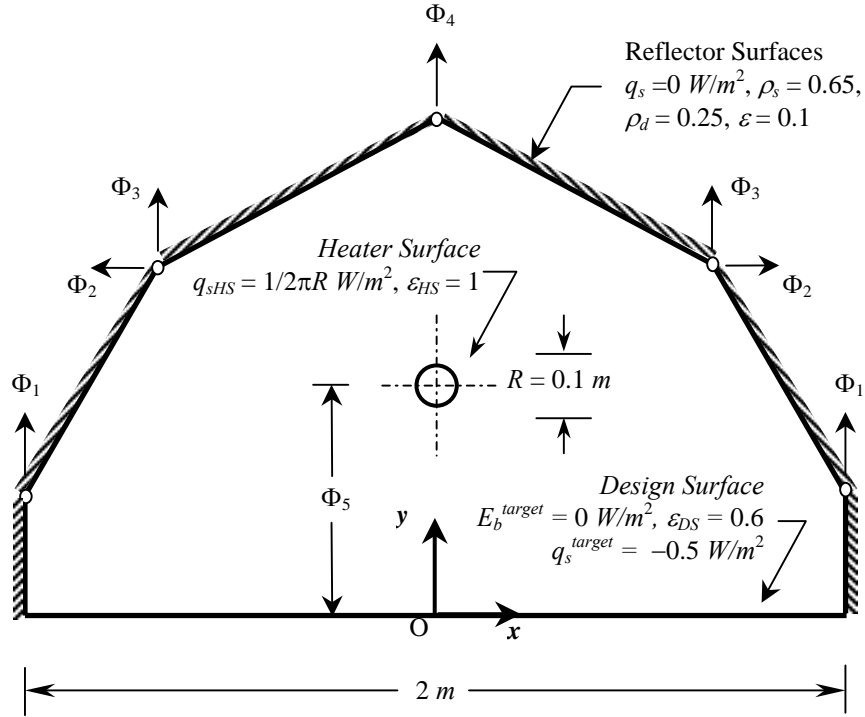


Figure 4.9: Radiant enclosure design problem involving multiple specular-diffuse surfaces.

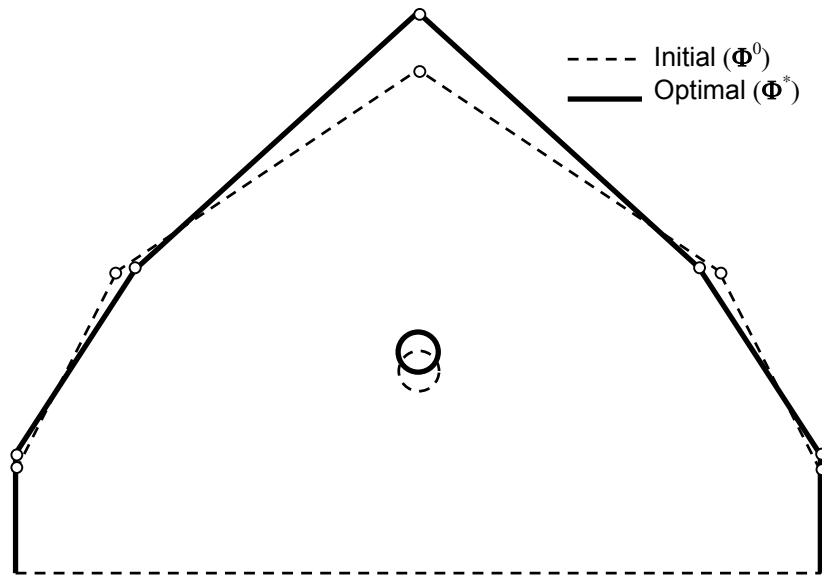


Figure 4.10: Enclosure geometries corresponding with  $\Phi^{0,1}$  and  $\Phi^{*,1}$ .

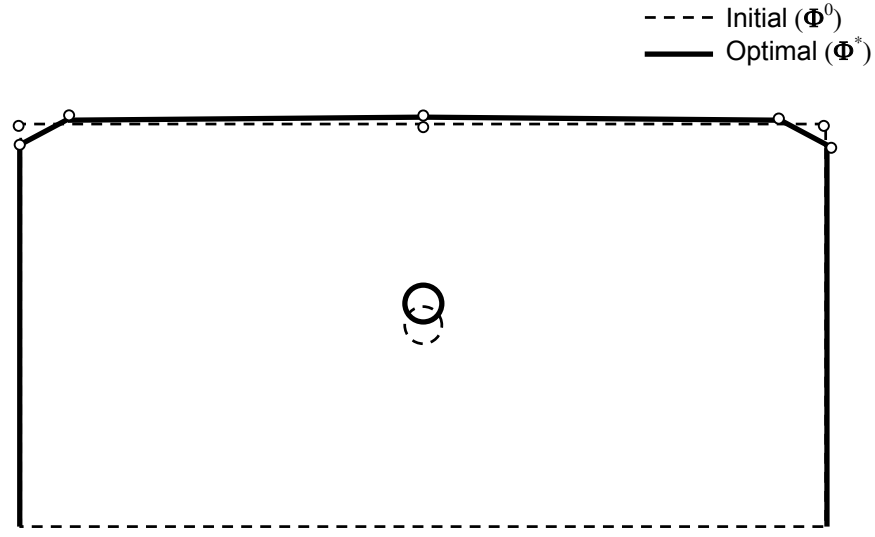


Figure 4.11: Enclosure geometries corresponding with  $\Phi^{0,2}$  and  $\Phi^{*,2}$ .

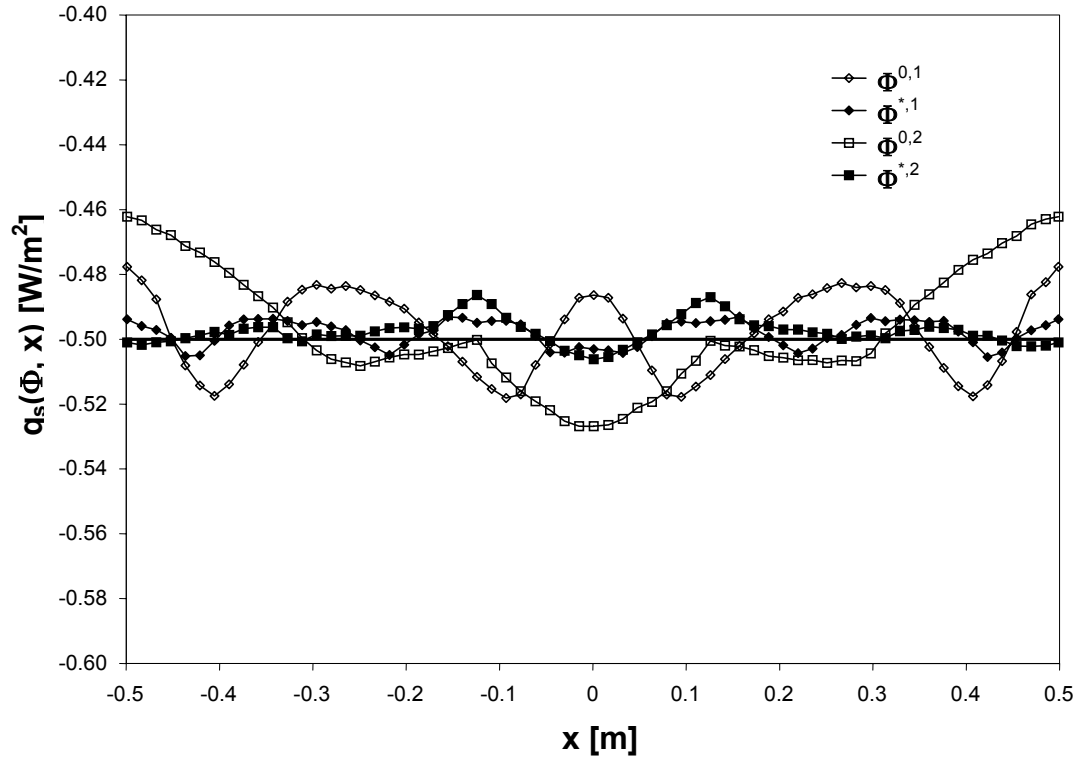


Figure 4.12: Heat flux distributions over the design surface produced by the initial and local optimal enclosure geometries.

## 4.5 Conclusions

This chapter presented an optimization methodology for designing the geometry and heater settings of radiant enclosures composed of non-diffuse surfaces. In particular, this methodology can be used to solve problems where directionally dependent surface properties must be modeled in order to obtain a sufficiently accurate solution, and also in cases where specularly-reflecting surfaces are used to obtain the desired conditions over the design surface.

While the problems presented in the previous chapters can be treated with a deterministic analysis, the complex nature of radiant exchange within these types of enclosure mandates the use of a Monte Carlo simulation to model heat transfer between the surfaces. The Monte Carlo technique was used to calculate the exchange factors between discrete surface elements, which in turn were used to solve for the heat flux distribution over the design surface. The objective function was minimized using the Kiefer-Wolfowitz method, which is specialized for optimizing stochastic systems and can accommodate the random error induced in the objective function by the Monte Carlo simulation.

## **Chapter 5: Optimization of Radiant Enclosures with Transient and Multimode Heat Transfer**



## 5.1 Introduction

The optimization methodologies presented thus far can only be used to solve design problems where the system is in steady state, and radiation is the only mode of heat transfer. Although the model of the physical system can sometimes be simplified to conform to these conditions without incurring unacceptable amounts of error, there are many problems where transient and multimode heat transfer effects cannot be ignored. This chapter presents an optimization methodology for solving this more complex class of radiant enclosure design problem.

Incorporating transient and multimode heat transfer effects into the analysis increases the number of problems that can be treated using the optimization design methodology. This is done, however, at the expense of creating a physical model that is more complicated and difficult to solve than those described in the previous chapters. While radiation heat transfer between enclosure surfaces is governed by integral equations involving absolute temperature to the fourth power, conjugate conduction, convection, and sensible energy storage are governed by differential equations involving temperature to the first power. In isolation, each of these effects can usually be modeled using linear systems of equations; combining them results in a highly nonlinear set of equations that is often difficult to solve.

Although this methodology can be adapted to solve a wide range of design problems, the focus here is on designing radiant enclosures that are part of a manufacturing process. In most cases, the enclosures are used to uniformly heat a product according to a desired temperature history or “heating curve”. Examples include

annealing furnaces used in foundries, baking ovens used in food preparation, infrared heating systems that cure painted surfaces, and rapid thermal processing (RTP) chambers used to manufacture semiconductor wafers. In each case, the radiant enclosure consists of a heater surface, several intermediate surfaces, and a design surface that contains the product to be processed. When designing these systems, it is important to have an accurate model of the transient heater settings in order to size the heaters and other components of the radiant enclosure. Having an accurate *a priori* estimate of the heater settings also enables the use of “high-gain” controllers under operating conditions, which quickly adjust the heaters to compensate for any deviation from the desired temperature distribution over the design surface.

In the past, the transient heater settings have been solved for using a forward “trial-and-error” technique. In this approach, the designer first guesses the appropriate heater settings, and then repeatedly evaluates and heuristically adjusts these settings until a satisfactory solution to the design problem is identified. As previously mentioned, forward design methodologies usually require many iterations and the final solution quality is limited; this is particularly true here, since the complicated nature of the coupled heat transfer modes make an intuitive understanding of the system physics elusive.

In order to overcome this difficulty, designers have adapted model-based control algorithms to design the transient heater settings. Control algorithms work as follows: at any time instant, the difference between the temperature measured at different locations on the design surface and the desired set-point temperature defines an error signal. This

error signal is passed through a feedback loop to the controller, which then adjusts the heater settings in such a way that the error signal is reduced. In most model-based control systems, the controller algorithm is formed from a set of first-order ordinary differential equations derived from a simplified heat transfer model; these equations relate the heater settings to the rate of temperature change over the design surface. Although they are most often used to control heaters under operating conditions, model-based control algorithms are also often used as design tools. For example, Breedijk et al. (1994) and Balakrishnan and Edgar (2002) both applied model-based controllers to design the heaters in RTP furnaces, and Yoshitani and Hasegawa (1998) designed a furnace used in a continuous annealing process. The solutions obtained using model-based control are usually found faster and better satisfy the desired temperature distributions than those found using the forward trial-and-error method.

Despite the fundamental non-linear nature of the problem, however, most model-based controllers use linear feedback algorithms that require the model to be composed of ordinary differential equations. Because of this, it is difficult to accommodate the integrals in the radiosity equation that account for reflection and reradiation, so these effects are usually ignored; this induces large modeling errors into the controller that severely limit the solution accuracy. (While the effects of modeling errors are mitigated through active measurement and feedback loops under operating conditions, there is no such mechanism in place during the design process.) In a recent technique presented by Gwak and Masada (2002), however, these effects are accounted for by applying non-

linear control laws coupled with embedded Tikhonov and TSVD regularization to solve the design problem.

More recently, several inverse design methodologies have been developed to solve this type of design problem. In this approach, both the desired temperature and the radiation heat input required to satisfy the sensible energy increase are specified over the design surface at different process times. The nonlinear system of equations is linearized, and the resulting set of ill-conditioned linear equations is solved using regularization methods, starting from the first step. Each solution is then used to define the right-hand-side vector of the next time step. França et al. (2001) were the first to propose an inverse design methodology for solving nonlinear radiant enclosure problems using TSVD regularization. This technique was demonstrated by solving a steady-state inverse boundary condition design problem consisting of an enclosure having heater and design surfaces located on opposing surfaces of a channel containing a hydrodynamically fully-developed flowing participating medium. The model took into account radiation exchange between the surfaces as well as convection and radiation exchange between the surfaces and the participating medium. A uniform number of singular values was truncated at each iteration in order to solve the ill-conditioned set of linear equations. Ertürk et al. (2002) later employed a similar technique to find the transient heater settings that heated a design surface moving through a roll-through batch furnace according to a prescribed temperature history. In this later work, a fixed number of conjugate gradient iterations was performed at each time step to regularize the solution of the ill-conditioned set of equations.

Unlike most model-based control algorithms, the inverse design approach can accommodate a very sophisticated heat transfer model resulting in a comparatively small modeling error. Nevertheless, a significant drawback of this method is that the error induced by regularization at a particular time step propagates throughout subsequent time steps, which limits the solution accuracy. Furthermore, as mentioned in previous chapters, it is very difficult to accommodate design constraints in the inverse design methodology, and because of this, the solutions found using regularization often include regions of negative heat flux over the heater surface. Since this condition cannot be realized in most practical furnaces, these regions are usually taken to be adiabatic, further impairing the solution quality.

Optimization through nonlinear programming overcomes many of these drawbacks. The general procedure is as follows: first, a suitable objective function,  $F(\Phi)$  is defined in such a way that the minimum of  $F(\Phi)$  corresponds to the desired design outcome, which in this case is a temperature distribution over the design surface that both matches the desired temperature history and is also uniform over the product throughout the process in order to ensure product homogeneity. The design parameters contained in  $\Phi$  define a set of functions that govern the heater output at any given time. Gradient-based minimization algorithms are then employed to find the set of design parameters,  $\Phi^*$ , that minimize the objective function, such that  $F(\Phi^*) = \text{Min}[F(\Phi)]$ . The design parameters contained in  $\Phi^*$  correspond to the transient heater settings that produce a temperature distribution over the design surface that most closely satisfies the design requirements.

Since the design parameters are modified in an intelligent way based on the local objective function curvature at each iteration, the optimization design methodology requires fewer iterations than the forward methodology, and the solution quality is much better. This technique can also accommodate a more sophisticated system model than most control algorithms and consequently is less susceptible to modeling errors. Finally, unlike the inverse design methodology, the optimization methodology can easily accommodate design constraints; in particular, it is convenient to enforce the heat flux generated over the heater surface to lie between zero and some maximum heater output value throughout the process.

Optimization techniques have been used on a limited basis to design industrial heating processes involving radiant enclosures. Norman (1992) used unconstrained linear programming to obtain the optimal heater settings for a simplified linearized model of an RTP furnace, and Cho and Gyugyi (1997) applied a similar procedure to obtain an initial estimate of the heater settings so that a high-gain controller could be used to operate an RTP furnace. Fedorov et al. (1998) used nonlinear programming to optimize the heater settings for a continuous roll-through industrial furnace operating under steady-state conditions.

This chapter presents two techniques for finding the optimal heater settings for two-dimensional, diffuse-walled radiant enclosures used in industrial heating processes. The first technique is generic in the sense that since sensible energy storage in the enclosure walls and conjugate conduction and convection effects are included in the heat transfer model, it can be used to solve design problems involving many different types of

radiant enclosures. The second technique is applied in cases where conduction and convection effects can be neglected, and where the design surface is the only enclosure surface that stores sensible energy. In both cases, a quasi-Newton minimization algorithm is again used to optimize the transient heater settings, and the design constraints are enforced through an active set method.

These optimization methodologies are demonstrated by applying them to solve industrially-relevant design problems. The first methodology is used to find the heater settings that heat a steel sheet according to a linear temperature history in a two-dimensional annealing furnace, while the second is used to solve for the heater settings that heat a silicon wafer according to a desired temperature history in an axisymmetric RTP furnace.

## 5.2 Optimization Methodology for Radiant Enclosures involving Conduction, Convection, and Transient Heat Transfer

This section presents an optimization methodology that is generic in the sense that it can be applied to solve a wide range of problems, since the physical model accommodates conduction heat transfer through the enclosure walls, convection heat transfer between the walls and surrounding fluids, and sensible energy storage in all the enclosure surfaces.

The first step of the optimization methodology is to define a suitable objective function and a set of design parameters that control the heat flux distribution over the heater surface throughout the process. Next, the integro-differential equations governing the temperature distribution and sensitivities are derived using an infinitesimal-area analysis. By following steps similar to those described in the previous chapters, the parametric region is discretized and the governing equations are transformed into nonlinear matrix equations, which are then solved. Finally, this technique is demonstrated by applying it to solve for the transient heater settings of a two-dimensional annealing furnace.

### 5.2.1 Optimization Strategy

Figure 5.1 shows an example radiant enclosure design problem that can be solved using this optimization methodology. In this example, the design surface is located on the bottom enclosure surface and is irradiated by heaters on the top surface, which in turn are controlled throughout the process by the design parameters contained in  $\Phi$ . The



underside of the design surface is insulated to prevent energy provided to product by the heaters from leaving the system. As is shown later, solving this problem requires the discretization of both the temporal and spatial (parametric) domains; the time domain is split into  $N_t$  time steps, the  $t^{th}$  time step having a duration  $\Delta t_t$  starting from time  $t_{t-1}$  and ending at  $t_t$ , and the design surface is split into  $N_{DS}$  discrete elements, with the  $i^{th}$  element having an area  $\Delta A_i$ . Furthermore, it is assumed that the density,  $\rho_{DS}$ , thermal capacity,  $c_{DS}$ , and thickness,  $\delta_{DS}$ , are uniform over the design surface, although this assumption can be discarded if necessary.

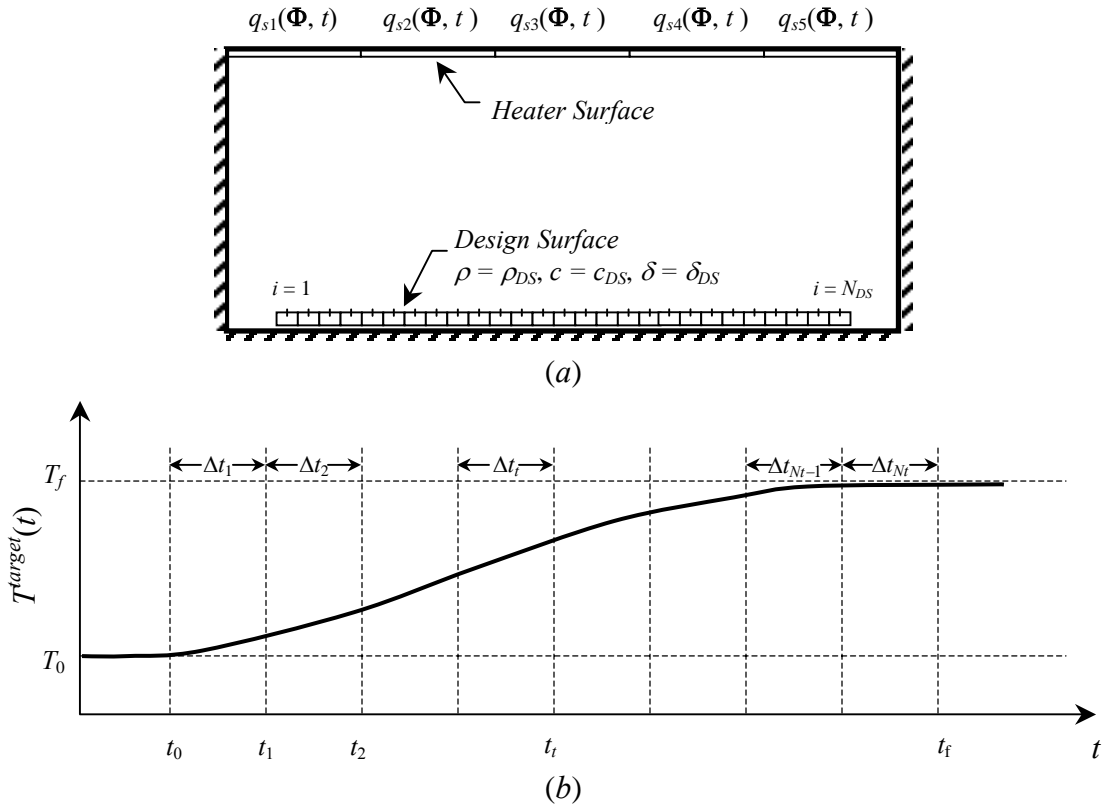


Figure 5.1: Example of a transient radiant enclosure design problem: (a) radiant enclosure, and (b) desired set-point temperature history.

The first step of this procedure is to define an objective function in such a way that it is minimized when (a) the average temperature over the design surface matches the desired set-point temperature at any instant, and (b) the temperature distribution over the design surface is uniform throughout the process. These design objectives could be satisfied individually by minimizing separate objective functions, each having a different minimum. Instead, these objective functions are combined to form a third objective function, whose minimum represents a trade-off between the two design objectives.

The first design objective is satisfied by minimizing the difference between the sensible energy provided to the design surface by the heaters, and the energy required to heat the design surface to the set-point temperature. The heater settings that accomplish this goal at each time step could be found by minimizing

$$F_1(\Phi) = \gamma_2 \sum_{t=1}^{N_t} [Q_{t-1 \rightarrow t}^{added}(\Phi) - Q_{t-1 \rightarrow t}^{target}(\Phi)]^2 + (1 - \gamma_2) [Q_{0 \rightarrow t_f}^{deficit}(\Phi)]^2, \quad (5.1)$$

where  $Q_{t-1 \rightarrow t}^{added}(\Phi)$  is the sensible energy increase in the design surface provided by the heaters during the  $t^{th}$  time step,

$$Q_{t-1 \rightarrow t}^{added}(\Phi) = \rho_{DS} c_{DS} \delta_{DS} \sum_{i=1}^{N_{DS}} [T_i(\Phi, t_t) - T_i(\Phi, t_{t-1})] \Delta A_i, \quad (5.2)$$

$Q_{t-1 \rightarrow t}^{target}$  is the sensible energy that must be added to achieve the desired temperature increase over the  $t^{th}$  time step,

$$Q_{t-1 \rightarrow t}^{target} = \rho_{DS} c_{DS} \delta_{DS} A_{DS} [T^{target}(t_t) - T^{target}(t_{t-1})], \quad (5.3)$$

and  $Q_{0 \rightarrow t_f}^{deficit}(\Phi)$  is the difference between the sensible energy added to the design surface throughout the process and what must be added to obtain the set point temperature at the end of the process,

$$Q_{0 \rightarrow t_f}^{deficit}(\Phi) = \rho_{DS} c_{DS} \delta_{DS} \sum_{i=1}^{N_{DS}} [T_i(\Phi, t_f) - T^{target}(t_f)], \quad (5.4)$$

These three quantities are shown schematically in Fig. 5.2. Minimizing the first part of  $F_1(\Phi)$  produces an average design surface temperature profile that best matches the slope of the set-point temperature. There is likely to be small difference between  $Q_{t-1 \rightarrow t}^{added}(\Phi^*)$  and  $Q_{t-1 \rightarrow t}^{target}$  at every time step, which accumulates over the process duration resulting in a net excess or deficit of sensible energy being added to the design surface. The second term in  $F_1(\Phi)$  ensures that this error is sufficiently small. The heuristic parameter  $\gamma_2$  is adjusted by the designer in order to achieve an average design surface temperature that closely matches the desired set-point temperature throughout the process.

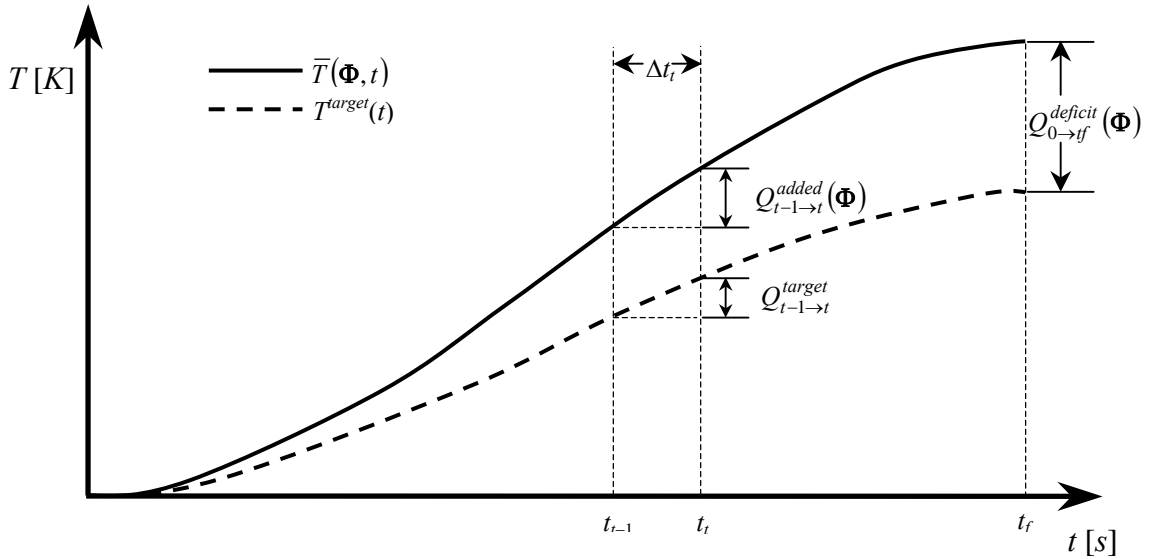


Figure 5.2: Relationship between  $Q_{t-1 \rightarrow t}^{added}(\Phi)$ ,  $Q_{t-1 \rightarrow t}^{target}$ ,  $Q_{0 \rightarrow t_f}^{deficit}(\Phi)$  and the design surface temperature.

The second design objective is to maintain a uniform temperature distribution over the design surface throughout the process. The heater settings that satisfy this

condition are found minimizing the variance between the temperature values calculated at discrete locations over the design surfaces, averaged over all time steps,

$$F_2(\Phi) = \frac{1}{N_t N_{DS}} \sum_{t=1}^{N_t} \sum_{i=1}^{N_{DS}} [T_i(\Phi, t_t) - \bar{T}(\Phi, t_t)]^2. \quad (5.5)$$

As previously mentioned, the objective functions defined in Eqs. (5.1) and (5.5) pertain to different design objectives, and both of these functions are minimized by different sets of  $\Phi$ , i.e.  $\Phi_1^* \neq \Phi_2^*$ . Since it is desirable to achieve both design objectives, a hybrid objective function is formed by combining these two objective functions,

$$F(\Phi) = C[\gamma_1 F_1(\Phi) + (1 - \gamma_1) F_2(\Phi)], \quad (5.6)$$

where  $\gamma_2$  is chosen based on the relative importance of the two design objectives in a particular application, and  $C$  is a scaling parameter. The set of design parameters  $\Phi^*$  that minimizes Eq. (5.6) is a compromise between a solution where the average design surface temperature closely follows the set-point temperature and one having a near-uniform temperature distribution over the design surface throughout the process.

Unlike the design methodologies presented in Chapters 2 and 3, the second-order design sensitivities needed to form the Hessian matrix are expensive to calculate, so the quasi-Newton's method is used in place of Newton's method to find  $\Phi^*$ . In the quasi-Newton's method, the search direction at the  $k^{th}$  iteration is found by solving

$$\mathbf{B}^k \mathbf{p}^k = -\nabla F(\Phi^k), \quad (5.7)$$

where  $\mathbf{B}^k$  approximates  $\nabla^2 F(\Phi)$  using first-order curvature information “built-up” over previous iterations. At the first iteration  $\mathbf{B}^0$  is set equal to the identity matrix and  $\mathbf{p}^0$  is the steepest-descent direction. The Hessian approximation is improved after each subsequent iteration by adding an update matrix,  $\mathbf{U}^k$ , which is formed using the first-order curvature

information collected from previous iterations. The update matrix is most often formed using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) scheme,

$$\mathbf{U}^k = -\frac{\mathbf{B}^k \mathbf{s}^k \mathbf{s}^{kT} \mathbf{B}^k}{\mathbf{s}^{kT} \mathbf{B}^k \mathbf{s}^k} + \frac{\mathbf{y}^k \mathbf{y}^{kT}}{\mathbf{y}^{kT} \mathbf{p}^k}, \quad (5.8)$$

with  $\mathbf{s}^k = \Phi^{k+1} - \Phi^k$  and  $\mathbf{y}^k = \nabla F(\Phi^{k+1}) - \nabla F(\Phi^k)$ .

Since  $\mathbf{B}^k$  accurately approximates  $\nabla^2 F(\Phi^k)$  and  $\mathbf{p}^k$  approaches Newton's direction only after a few iterations, the quasi-Newton's method usually requires more iterations than Newton's method to find  $\Phi^*$ . Because of the computational effort and storage required at each iteration to calculate  $\nabla^2 F(\Phi^k)$ , however, the quasi-Newton's method is more efficient than Newton's method in this particular application.

From examining Eqs. (5.7) and (5.8), it is clear that the quasi-Newton method requires only the first-order objective function sensitivities to calculate the search direction; these are found by differentiating  $F(\Phi)$  with respect to each design parameter,

$$\mathbf{g}_p(\Phi) = \frac{\partial F(\Phi)}{\partial \Phi_p} = C \left[ \gamma_1 \frac{\partial F_1(\Phi)}{\partial \Phi_p} + (1 - \gamma_1) \frac{\partial F_2(\Phi)}{\partial \Phi_p} \right]. \quad (5.9)$$

The sensitivities of  $F_1(\Phi)$  in turn are calculated by

$$\frac{\partial F_1(\Phi)}{\partial \Phi_p} = 2 \sum_{t=1}^{N_t} Q_{t-1 \rightarrow t}^{added}(\Phi) \frac{\partial Q_{t-1 \rightarrow t}^{added}(\Phi)}{\partial \Phi_p} + 2 Q_{0 \rightarrow t_f}^{deficit}(\Phi) \frac{\partial Q_{0 \rightarrow t_f}^{deficit}(\Phi)}{\partial \Phi_p}, \quad (5.10)$$

with

$$\frac{\partial Q_{t-1 \rightarrow t}^{added}(\Phi)}{\partial \Phi_p} = \rho_{DS} c_{DS} \delta_{DS} \sum_{i=1}^{N_{DS}} \left[ \frac{\partial T_i(\Phi, t_t)}{\partial \Phi_p} - \frac{\partial T_i(\Phi, t_{t-1})}{\partial \Phi_p} \right] \Delta A_i, \quad (5.11)$$

and

$$\frac{\partial Q_{0 \rightarrow t_f}^{deficit}(\Phi)}{\partial \Phi_p} = \rho_{DS} c_{DS} \delta_{DS} \sum_{i=1}^{N_{DS}} \frac{\partial T_i(\Phi, t_f)}{\Phi_p} \Delta A_i, \quad (5.12)$$

while the sensitivities of  $F_2(\Phi)$  are given by

$$\frac{\partial F_2(\Phi)}{\partial \Phi_p} = \frac{2}{N_t N_{DS}} \sum_{t=1}^{N_t} \sum_{i=1}^{N_{DS}} [T_i(\Phi, t_t) - \bar{T}(\Phi, t_t)] \left[ \frac{\partial T_i(\Phi, t_t)}{\partial \Phi_p} - \frac{\partial \bar{T}(\Phi, t_t)}{\partial \Phi_p} \right], \quad (5.13)$$

with

$$\frac{\partial \bar{T}(\Phi, t_t)}{\partial \Phi_p} = \frac{1}{N_{DS}} \sum_{i=1}^{N_{DS}} \frac{\partial T_i(\Phi, t_t)}{\partial \Phi_p}. \quad (5.14)$$

Techniques for solving the temperature distribution and temperature sensitivities needed to find  $F(\Phi)$  and  $\partial F(\Phi)/\partial \Phi_p$  are presented in the next section.

### 5.2.2 Calculation of the Temperature Distribution and Sensitivities

An infinitesimal-area analysis technique similar to the one used in Chapters 2 and 3 is used to derive the equations governing the temperature and temperature sensitivities. The first step of the analysis is to identify a suitable parametric representation for the enclosure. As before, the enclosure geometry is specified by

$$\mathbf{r} = \mathbf{C}(u) = \{P(u), Q(u)\}^T, \quad a \leq u \leq b, \quad (5.15)$$

and either the temperature,  $T(u, \Phi, t)$ , or the heat flux,  $q_s(u, \Phi, t)$ , is assumed to be known at every location on the enclosure surface at any time  $t$ . In particular, the transient heat flux distribution over the heater surfaces is specified as a function of the heater settings contained in  $\Phi$  and the adiabatic boundary condition is enforced over the design surface throughout the process. The emissivity,  $\varepsilon(u)$ , thermal conductivity,  $\kappa(u)$ , density,  $\rho(u)$ , specific heat,  $c(u)$ , and wall thickness,  $\delta(u)$ , of the enclosure walls are also specified parametrically.

Once parametric representation of the enclosure has been identified, the equation relating the radiosity distribution,  $q_o(u, \Phi, t)$ , to the temperature distribution,  $T(u, \Phi, t)$ , is derived by performing an energy balance on an infinitely long wall element having a thickness  $\delta(u)$  and an infinitesimal chord length  $J(u)du$ , as shown in Fig. 5.3, where the surface discriminant,  $J(u)$ , is given by

$$J(u) = \left\{ \left[ \frac{\partial P(u)}{\partial u} \right]^2 + \left[ \frac{\partial Q(u)}{\partial u} \right]^2 \right\}^{1/2}. \quad (5.16)$$

By setting the net rate of conduction, convection, and thermal radiation heat transfer entering the infinitesimal wall element equal to the rate of sensible energy storage, it can be shown that the radiosity and temperature distributions are related by

$$\begin{aligned}
& -q_o(u, \Phi, t) + \int_a^b q_o(u', \Phi, t) k(u, u') du' + q_s(u, \Phi, t) \\
& + q_{cond}(u, \Phi, t) - q_{conv}(u, \Phi, t) = \rho(u) c(u) \delta(u) \frac{\partial T(u, \Phi, t)}{\partial t}, \quad (5.17)
\end{aligned}$$

where  $k(u, u)$  is the kernel defined in Eq. (3.7). Besides radiant heat transfer, there are three other modes of heat transfer entering and leaving the wall element:  $q_{cond}(u, \Phi, t)$  is the net rate of heat transfer entering the wall element by conduction from the surrounding enclosure wall,

$$q_{cond}(u, \Phi, t) = \frac{1}{J(u)} \frac{\partial}{\partial u} \left[ \frac{\kappa(u)}{J(u)} \frac{\partial T(u, \Phi, t)}{\partial u} \right], \quad (5.18)$$

$q_{conv}(u, \Phi, t)$  is the rate of convection transferred from the wall element to the fluid contained within the enclosure,

$$q_{conv}(u, \Phi, t) = h(u, t) [T(u, \Phi, t) - T_\infty(t)], \quad (5.19)$$

and  $q_s(u, \Phi, t)$  is the rate that any other type of heat transfer enters the wall element by non-radiative means; all of these terms are per unit area of the infinitesimal wall element on the interior enclosure surface. (Convection heat transfer with the fluid surrounding the enclosure has been excluded to simplify the heat transfer model, but could be added if necessary.)

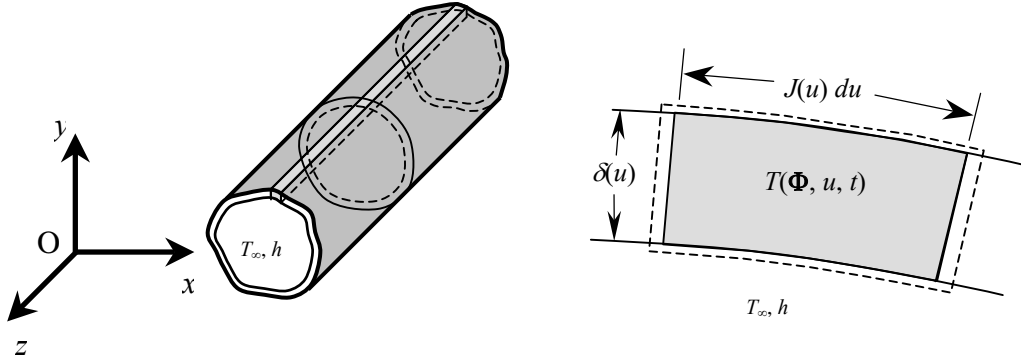


Figure 5.3: Infinitesimal wall element used to form governing temperature equation.



Equation (5.17) cannot be solved directly, however, because both the temperature and the radiosity distributions are unknown. A more convenient form is found by performing another energy balance on the infinitesimal wall element at  $u'$ ; this results in a non-linear integro-differential equation where temperature is the only variable,

$$\sigma T^4(u, \Phi, t) - \int_a^b \sigma T^4(u', \Phi, t) k(u, u') du' = \frac{b(u, \Phi, t)}{\varepsilon(u)} - \int_a^n \left[ \frac{1 - \varepsilon(u')}{\varepsilon(u')} \right] b(u', \Phi, t) k(u, u') du', \quad (5.20)$$

where  $b(u, \Phi, t)$  represents the difference between the net non-radiative heat transfer into an infinitesimal wall element and the sensible energy stored within that element,

$$b(u, \Phi, t) = q_s(u, \Phi, t) + q_{cond}(u, \Phi, t) - q_{conv}(u, \Phi, t) - \rho(u)c(u)\delta(u)\frac{\partial T(u, \Phi, t)}{\partial t}. \quad (5.21)$$

The equations governing the temperature sensitivity are found by differentiating Eqs. (5.20) and (5.21) with respect to the design parameter of interest. By applying Liebnitz's rule to Eq. (5.20) and noting that the integral bounds are independent of  $u$ , it is shown that the temperature sensitivities are governed by the related integro-differential equations

$$4\sigma T^3(u, \Phi, t) \frac{\partial T(u, \Phi, t)}{\partial \Phi_p} - \int_a^b 4\sigma T^3(u', \Phi, t) k(u, u') du' = \frac{1}{\varepsilon(u)} \frac{\partial b(u, \Phi, t)}{\partial \Phi_p} - \int_a^n \left[ \frac{1 - \varepsilon(u')}{\varepsilon(u')} \right] \frac{\partial b(u', \Phi, t)}{\partial \Phi_p} k(u, u') du', \quad (5.22)$$

where

$$\begin{aligned} \frac{\partial b(u, \Phi, t)}{\partial \Phi_p} = & \frac{\partial q_s(u, \Phi, t)}{\partial \Phi_p} + \frac{1}{J(u)} \frac{\partial}{\partial u} \left[ \frac{\kappa(u)}{J(u)} \frac{\partial^2 T(u, \Phi, t)}{\partial u \partial \Phi_p} \right] \\ & - h(u, t) \frac{\partial T(u, \Phi, t)}{\partial \Phi_p} - \rho(u)c(u)\delta(u) \frac{\partial^2 T(u, \Phi, t)}{\partial t \partial \Phi_p}. \end{aligned} \quad (5.23)$$

Since analytical solutions to integro-differential equations are usually not tractable, the temperature and temperature sensitivity distributions must be solved numerically. The first step is to discretize the parametric and temporal domains as described in the previous section. The parametric domain is divided into  $N$  elements, with the  $i^{th}$  element centered on  $u_i$  and having a width  $\Delta u_i$ . Each of the elements in parametric space corresponds to an infinitely long wall element having a finite thickness, as shown in Fig. 5.4. The time domain is discretized into  $N_t$  time steps starting from  $t_0$  to  $t_{N_t} = t_f$  proceeding in intervals of  $\Delta t_t$ .

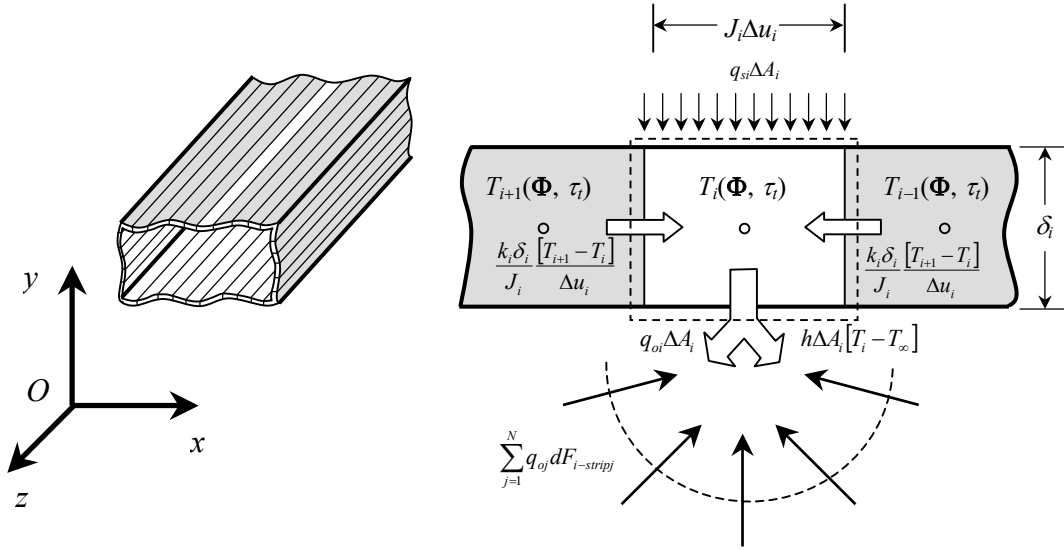


Figure 5.4: Discretization of the radiant enclosure into finite wall elements.

Once this is done, the integrals and derivatives with respect to  $t$  and  $u$  in the governing equations are represented with discrete approximations. First, the integrals in

Eq. (5.20) and (5.22) are approximated using a method analogous to the one described in Chapters 2 and 3, so that

$$\int_a^b x(\Phi, u', t) k(u, u') du' \approx \sum_{j=1}^N x(\Phi, u_j, t) dF_{i-stripj}, \quad (5.24)$$

where  $x(u, \Phi, t)$  is the integrated quantity and  $dF_{i-stripj}$  is the configuration factor between a point on the enclosure surface at  $u_i$  and the exposed surface of an infinitely long wall element centered at  $u_j$ , defined in Eq. (3.20).

Next, the time derivatives in Eq. (5.21) are approximated using a first-order backwards operator,

$$\frac{\partial T(\Phi, u_i, t)}{\partial t} \approx \frac{T(\Phi, u_i, t) - T(\Phi, u_i, t_{t-1})}{\Delta t_t}, \quad (5.25)$$

with similar expressions applying to the time derivatives in Eq. (5.23). Finally, the derivatives in Eq. (5.18) associated with thermal conduction are approximated using a second-order finite difference scheme. If  $\kappa(u)$  and  $J(u)$  are uniform over each enclosure surface (which is often the case) and if a uniform quadrature is used over the parametric sub-domain corresponding to each enclosure surface, then

$$q_{cond}(\Phi, u_i, t) \approx \frac{\kappa(u_i)}{J(u_i)} \left[ \frac{T(\Phi, u_{i+1}, t) - 2T(\Phi, u_i, t) + T(\Phi, u_{i-1}, t)}{2\Delta u_i} \right]. \quad (5.26)$$

(This is assumed throughout the remainder of this section.) Similar relations apply to the sensitivity of the thermal conductivity with respect to the design parameters in Eq. (5.23).

Substituting the above approximations into the integro-differential equation governing the temperature distribution and rewriting the parameters in non-dimensional form results in

$$\hat{T}_i^4(\Phi, \tau_t) - \sum_{\substack{j=1 \\ j \neq i}}^N \hat{T}_j^4(\Phi, \tau_t) dF_{i-stripj} = \frac{\hat{b}_i(\Phi, \tau_t, \tau_{t-1})}{\varepsilon_i} - \sum_{\substack{j=1 \\ j \neq i}}^N \left[ \frac{1 - \varepsilon_j}{\varepsilon_j} \right] \hat{b}_j(\Phi, \tau_t, \tau_{t-1}) dF_{i-stripj}, \quad (5.27)$$

where  $\varepsilon_i = \varepsilon(u_i)$ , and the time and temperature terms are represented non-dimensionally by  $\tau_t = (t_t - t_0)/(t_f - t_0)$ , and  $\hat{T}_i(\Phi, \tau_t) = T(\Phi, u_i, t_t)/T_s$ , where  $T_s$  is a scaling temperature. The term  $\hat{b}_i(\Phi, \tau_t, \tau_{t-1})$  is the difference between the net rate of non-radiative heat transfer to the  $i^{th}$  element and the rate of sensible energy storage by the element during the  $t^{th}$  time step,

$$\begin{aligned} \hat{b}_i(\Phi, \tau_t, \tau_{t-1}) = & \hat{q}_{si}(\Phi, \tau_t) + C_{cond} \frac{\hat{\kappa}_i}{\hat{J}_i^2} \left[ \frac{\hat{T}_{i+1}(\Phi, \tau_t) - 2\hat{T}_i(\Phi, \tau_t) + \hat{T}_{i-1}(\Phi, \tau_t)}{2\Delta u_i} \right] \\ & + C_{conv} \hat{h}_i(\tau_t) [\hat{T}_i(\Phi, \tau_t) - \hat{T}_\infty(\tau_t)] + C_{trans} \hat{\rho}_i \hat{c}_i \hat{\delta}_i \frac{\hat{T}_i(\Phi, \tau_t) - \hat{T}_i(\Phi, \tau_{t-1})}{\Delta \tau_t}, \end{aligned} \quad (5.28)$$

where  $\hat{q}_{si}(\Phi, \tau_t) = q_s(\Phi, u_i, t_t)/\sigma T_s^4$  and  $\hat{h}_i(\tau_t) = h(u_i, t_t)/\bar{h}$ , with  $\bar{h}$  being the average heat transfer coefficient over the enclosure surfaces. The enclosure wall properties are represented non-dimensionally by  $\hat{\kappa}_i = \kappa(u_i)/\kappa_s$ ,  $\hat{\rho}_i = \rho(u_i)/\rho_s$ ,  $\hat{c}_i = c(u_i)/c_s$ ,  $\hat{\delta}_i = \delta(u_i)/\delta_s$  and  $\hat{J}_i(\tau_k) = J(u_i, t_k)/L_c$ , where  $L_c$  is a characteristic length. The non-dimensional coefficients  $C_{cond}$ ,  $C_{conv}$ , and  $C_{trans}$  are

$$C_{cond} = \frac{\kappa_s}{\sigma T_s^3 L_c^2}, \quad (5.29)$$

$$C_{conv} = \frac{\bar{h}}{\sigma T_s^3}, \quad (5.30)$$

and

$$C_{trans} = \frac{\rho_s c_s \delta_s}{\sigma T_s^3 (t_f - t_0)}, \quad (5.31)$$

respectively, and their magnitudes signify the relative importance of conduction, convection and sensible energy storage compared to thermal radiation in the heat transfer model.

Writing Eqs. (5.27) and (5.28) for every element results in a nonlinear system of  $N$  equations containing  $N$  unknowns that can be rewritten as a matrix equation governing the temperature distribution over the enclosure surface at time  $\tau_i$ ,

$$\mathbf{A}_1 \mathbf{x}_1(\Phi, \tau_i) + \mathbf{A}_2 \mathbf{x}_2(\Phi, \tau_i) = \mathbf{c}(\Phi, \tau_{i-1}), \quad (5.32)$$

where  $x_{1,i}(\Phi, \tau_i) = \hat{T}_i^4(\Phi, \tau_i)$  and  $x_{2,i}(\Phi, \tau_i) = \hat{T}_i(\Phi, \tau_i)$ . The transient temperature distribution over the enclosure surface could be found by first guessing an initial temperature distribution at  $\tau = 0$  and then writing and solving Eq. (5.32) at each subsequent time step using the temperature distribution from the previous time step to form  $\mathbf{c}(\Phi, \tau_{i-1})$  at the current time step. More often, however, the temperature distribution is solved by linearizing Eq. (5.32) to form a related matrix equation

$$\mathbf{A}(\Phi, \tau_{i-1}) \mathbf{x}(\Phi, \tau_i) = \mathbf{b}(\Phi, \tau_{i-1}), \quad (5.33)$$

where  $x_i(\Phi, \tau_i)$  is either  $\hat{T}_i^4(\Phi, \tau_i)$  or  $\hat{T}_i(\Phi, \tau_i)$ , depending on how the problem is linearized. Since not all linearization schemes will produce a convergent solution for a given problem, the method used to linearize Eq. (5.32) must be carefully chosen based on the relative magnitudes of the coefficients defined in Eqs. (5.29) – (5.31). Siegel and Howell (2002b) provide a detailed discussion of these linearization methods, and one is demonstrated in the next section.

The temperature sensitivities are calculated in a similar manner. The integrals and derivatives in Eqs. (5.22) and (5.23) are first represented discretely as described in Eqs. (5.24) – (5.26), and the integro-differential equation governing the temperature sensitivity distribution over the enclosure surface is then rewritten in discrete form as

$$\begin{aligned} 4\hat{T}_i^3(\Phi, \tau_t) \frac{\partial \hat{T}_i(\Phi, \tau_t)}{\partial \Phi_p} - \sum_{\substack{j=1 \\ j \neq i}}^N 4\hat{T}_j^3(\Phi, \tau_t) \frac{\partial \hat{T}_j(\Phi, \tau_t)}{\partial \Phi_p} dF_{i-stripj} \\ = \frac{1}{\varepsilon_i} \frac{\partial \hat{b}_i(\Phi, \tau_t, \tau_{t-1})}{\partial \Phi_p} - \sum_{\substack{j=1 \\ j \neq i}}^N \left( \frac{1 - \varepsilon_j}{\varepsilon_j} \right) \frac{\partial \hat{b}_j(\Phi, \tau_t, \tau_{t-1})}{\partial \Phi_p} dF_{i-stripj}, \end{aligned} \quad (5.34)$$

where

$$\begin{aligned} \frac{\partial \hat{b}_i(\Phi, \tau_t, \tau_{t-1})}{\partial \Phi_p} = \frac{\partial \hat{q}_s(\Phi, \tau_t)}{\partial \Phi_p} + C_{cond} \frac{\kappa_i}{\Delta u_i^2} \left[ \frac{\partial T_{i+1}(\Phi, \tau_t)}{\partial \Phi_p} - 2 \frac{\partial T_i(\Phi, \tau_t)}{\partial \Phi_p} + \frac{\partial T_{i-1}(\Phi, \tau_t)}{\partial \Phi_p} \right] \\ - C_{conv} \hat{h}_i(\tau_t) \frac{\partial T_i(\Phi, \tau_t)}{\partial \Phi_p} - C_{trans} \frac{\hat{\rho}_i \hat{c}_i \hat{\delta}_i}{\Delta \tau_t} \left[ \frac{\partial T_i(\Phi, \tau_t)}{\partial \Phi_p} - \frac{\partial T_i(\Phi, \tau_{t-1})}{\partial \Phi_p} \right]. \end{aligned} \quad (5.35)$$

If the temperature distribution has already been solved for, writing Eqs. (5.34) and (5.35) for every wall element results in a system of  $N$  equations containing  $N$  unknowns that can be rearranged to form a matrix equation,

$$\mathbf{dA}(\Phi, \tau_{t-1}) \mathbf{x}'(\Phi, \tau_t) = \mathbf{b}'(\Phi, \tau_{t-1}), \quad (5.36)$$

where  $x'_i(\Phi, \tau_t) = \partial \hat{T}_i(\Phi, \tau_t) / \partial \Phi_p$ . (This coefficient matrix is dependant on the design parameters because the non-dimensional temperatures are coefficients of the temperature sensitivities in Eq. (5.33).)

As with the temperature distribution, the temperature sensitivities are solved by first guessing a sensitivity distribution at  $\tau = 0$ , and then writing and solving Eq. (5.36) at each time step using the temperature sensitivities from the previous time step to form the

right-hand vector corresponding to the current time step. Although Eq. (5.36) is linear and can be solved directly at each time step, it is often more convenient to use a modified form consistent with the linearization scheme used to find the temperature distribution, since the  $\mathbf{A}$  matrix in Eq. (5.33) and the  $\mathbf{dA}$  matrix in Eq. (5.36) are sometimes identical. If this is the case, the temperature sensitivities can be easily solved once the  $\mathbf{A}$  matrix has been decomposed to find the temperature distribution. This is demonstrated in the following example.

### 5.2.3 Implementation and Demonstration

The optimization design methodology described in the previous section is demonstrated by applying it to solve for the heater settings in the two-dimensional annealing furnace shown in Fig. 5.5. The annealing furnace is 1 *m* wide and 0.5 *m* high; the top surface contains ten uniformly-spaced heaters, the two side walls are refractory surfaces, and the design surface is on the bottom of the enclosure. The heater and refractory surfaces are assumed to have the properties of 10 *cm* thick refractory brick, while the design surface is a 2 *cm* thick slab of AISI 1010 steel. The enclosure wall properties are summarized in Table 5.1.

The objective of this problem is to uniformly heat the steel according to a linear ramp rate from an initial temperature of  $T_0 = 300\text{ K}$  to a final temperature of  $T_f = 500\text{ K}$ , over a duration of 5 hours. It is assumed that the enclosure surfaces are initially at thermal equilibrium at 300 *K*. The heaters are activated at  $t = 0$ , and all the surfaces are exposed to a fluid at  $T_\infty = 500\text{ K}$  with a convection coefficient  $h = 5\text{ W/m}^2\text{K}$  between the

enclosure surfaces and the fluid throughout the process. The enclosure surfaces are assumed to be thermally isolated from each other, which is enforced by insulating the surface edges.

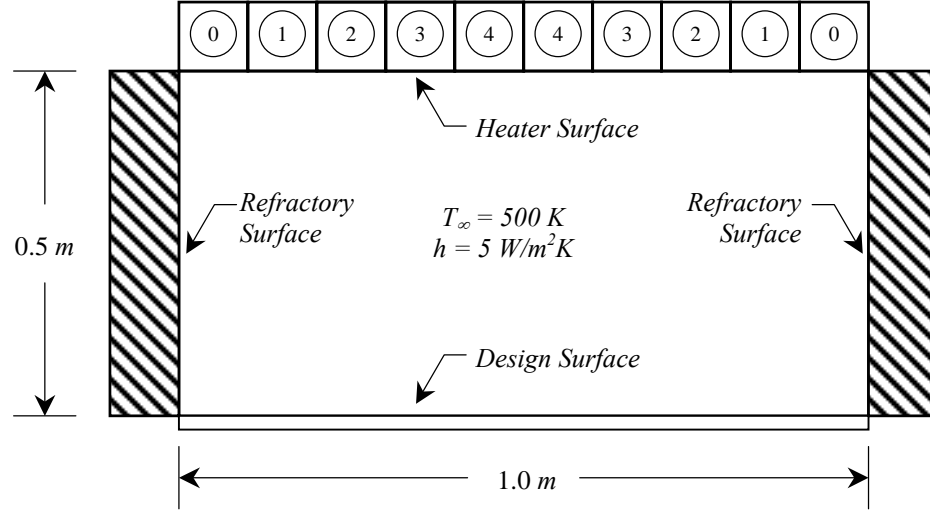


Figure 5.5: Example design problem involving conduction, convection, and transient heat transfer effects. (Heater numbers are shown in circles.)

	$\kappa [\text{W/mK}]$	$\rho [\text{kg/m}^3]$	$c [\text{J/kgK}]$	$\delta [\text{m}]$	$\varepsilon$
Heater Surface	1.0	2645	960	0.1	0.8
Refractory Surface	1.0	2645	960	0.1	0.8
Design Surface	63.9	7832	487	0.02	0.4

Table 5.1: Surface properties corresponding to the enclosure of Fig. 5.4.

The heat flux distribution over each heater is assumed to be uniform and is controlled by a cubic spline function of non-dimensional time. Because of the problem symmetry, the heaters are controlled in pairs and are numbered as shown in Fig. 5.5. In particular, if  $u_i$  lies on the  $h^{\text{th}}$  heater,



$$\hat{q}_{si}(\Phi, \tau) = \Phi_{4h}(1-\tau)^3 + 3\Phi_{4h+1}\tau(1-\tau)^2 + 3\Phi_{4h+2}\tau^2(1-\tau) + \Phi_{4h+3}\tau^3, \quad (5.37)$$

where  $\{\Phi_{4h}, \Phi_{4h+1}, \Phi_{4h+2}, \Phi_{4h+3}\}^T$  is a vector subspace of  $\Phi$ ; accordingly, 20 design parameters are required to specify the heat flux distribution over the heater surface throughout the process. Controlling the heater output in this way reduces the dimension of the minimization problem (versus optimizing the heater settings at each time step) and acts to regularize the solution since cubic splines are smooth functions. Also, since the cubic basis functions in Eq. (5.37) sum to unity for any value of  $\tau$  throughout the process, the heater outputs can be constrained to lie within a specific operating range,  $l_{bh} \leq q_{si}(\Phi, \tau) \leq u_{bh}$ , by using these values to define the lower and upper bounds of the design parameters that control the  $h^{th}$  heater. In this problem, the non-dimensional heat flux is constrained to lie in the range defined by  $0 \leq \hat{q}_{si}(\Phi, \tau) \leq 10$ , which is enforced by incorporating an active set method into the BFGS minimization. (This technique is described in detail in Appendix A.)

The problem is non-dimensionalized using  $L_c = 1 \text{ m}$ ,  $\bar{h} = 5 \text{ W/m}^2\text{K}$ , and  $T_s = 1000 \text{ K}$ , which is approximately the maximum temperature on the heater surface reached during the process, while  $\kappa_s$ ,  $\rho_s$ ,  $c_s$ , and  $\delta_s$  are set equal to the design surface properties. Substituting these values into the coefficients defined in Eqs. (5.29) – (5.31) results in  $C_{cond} = 1.1270$ ,  $C_{conv} = 0.0882$ , and  $C_{trans} = 0.0747$ .

Examining the magnitudes of the non-dimensional coefficients indicates that conduction effects dominate the heat transfer model. Accordingly, Eq. (5.27) is linearized by lagging the emissive power terms,

$$\begin{aligned}
& \hat{T}_i^4(\Phi, \tau_{t-1}) - \sum_{\substack{j=1 \\ j \neq i}}^N \hat{T}_j^4(\Phi, \tau_{t-1}) dF_{i-stripj} \\
& = \frac{\hat{b}_i(\Phi, \tau_t, \tau_{t-1})}{\varepsilon_i} - \sum_{\substack{j=1 \\ j \neq i}}^N \left[ \frac{1 - \varepsilon_j}{\varepsilon_j} \right] \hat{b}_i(\Phi, \tau_t, \tau_{t-1}) dF_{i-stripj}, \quad (5.38)
\end{aligned}$$

where  $\hat{b}_i(\Phi, \tau_t, \tau_{t-1})$  is defined in Eq. (5.28). The adiabatic boundary conditions at the surface edges are enforced using second-order forward and backward finite difference schemes.

Writing (5.38) for each wall element results in a linear system of  $N$  equations containing  $N$  unknowns that can be rearranged into the matrix equation

$$\mathbf{A} \mathbf{x}(\Phi, \tau_t) = \mathbf{b}(\Phi, \tau_{t-1}), \quad (5.39)$$

where  $\mathbf{A}$  contains the conduction, convection, and sensible energy storage temperature coefficients,  $\mathbf{b}(\Phi, \tau_{t-1})$  is composed of terms associated with the conduction boundary condition, fluid temperature, and sensible energy and thermal radiation terms from the previous time steps, and  $\mathbf{x}(\Phi, \tau_t) = \{\hat{T}_1(\Phi, \tau_t), \hat{T}_2(\Phi, \tau_t), \dots, \hat{T}_N(\Phi, \tau_t)\}^T$ . If this linearization scheme is also applied to Eqs. (5.34) and (5.35), it can be shown that the temperature sensitivities are governed by

$$\mathbf{A} \mathbf{x}'(\Phi, \tau_t) = \mathbf{b}'(\Phi, \tau_{t-1}), \quad (5.40)$$

where  $\mathbf{b}'(\Phi, \tau_{t-1})$  contains the sensitivities of the sensible energy and thermal radiation terms with respect to  $\Phi_p$ , and  $\mathbf{x}'_i(\Phi, \tau_t) = \partial \hat{T}_i(\Phi, \tau_t) / \partial \Phi_p$ . Thus,  $\mathbf{A}$  needs to be formed and decomposed only once at each time step in order to calculate both the temperature and the complete set of temperature sensitivities.

The optimal heater settings are found by minimizing the objective function defined in Eq. (5.6). The parametric and time domains were discretized uniformly, and a total of  $N = 240$  wall elements and  $N_t = 500$  time steps were used to calculate  $F(\Phi)$  and  $\nabla F(\Phi)$  throughout the optimization process. The minimization was carried out starting from  $\Phi_i^0 = 1$ ,  $i = 1 \dots 20$ , and was stopped when  $\|\nabla F_{FR}(\Phi^k)\|_2 \leq 10^{-6}$ , where  $\nabla F_{FR}(\Phi^k)$ , contains the first-order objective function sensitivities with respect to the unconstrained design parameters. After trying several different values for  $C$ ,  $\gamma_1$ , and  $\gamma_2$ , a good result was finally obtained with  $C = 100$ ,  $\gamma_1 = 0.99$  and  $\gamma_2 = 0.995$ .

A local minimum of  $F(\Phi^*) = 2.284 \times 10^{-3}$  was found after 30 iterations. The optimal heater settings are shown in Fig. 5.6, while the resulting transient temperature distribution over the design surface is shown in Figs. 5.7 and 5.8. Figure 5.7 shows that the design surface temperature was less than the set-point temperature for  $0 < \tau < 0.472$ , and exceeded the set-point temperature for the remainder of the process; the maximum deviation of the average temperature from the set point temperature was 7.15% at the end of the process. While a better solution might be found by using higher-order splines to control the heaters, the thermal inertia of the enclosure walls severely limits the response sensitivity of the design surface temperature to the design parameters. In particular, a significant amount of the overshoot error is likely due to heat transfer between the design surface and the hot refractory walls.

As shown in Fig. 5.8, a near-uniform temperature distribution is maintained throughout the process; the maximum standard deviation from the average design surface temperature is less than  $\sigma = 0.006 \%$ , occurring at  $\tau = 1$ .

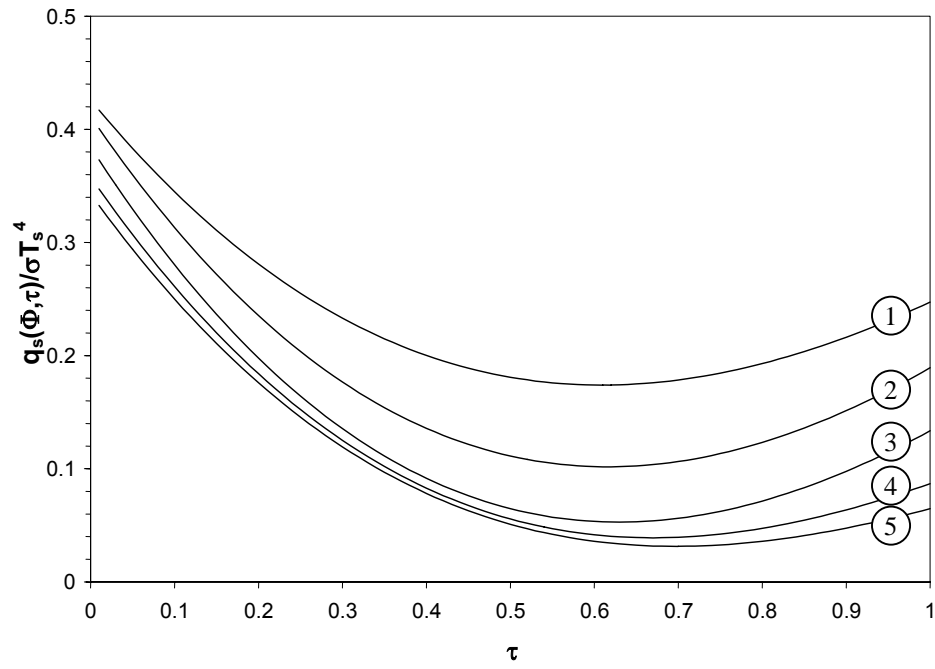


Figure 5.6: Optimal heater settings for the problem shown in Fig. 5.5. (Heater numbers shown in circles.)

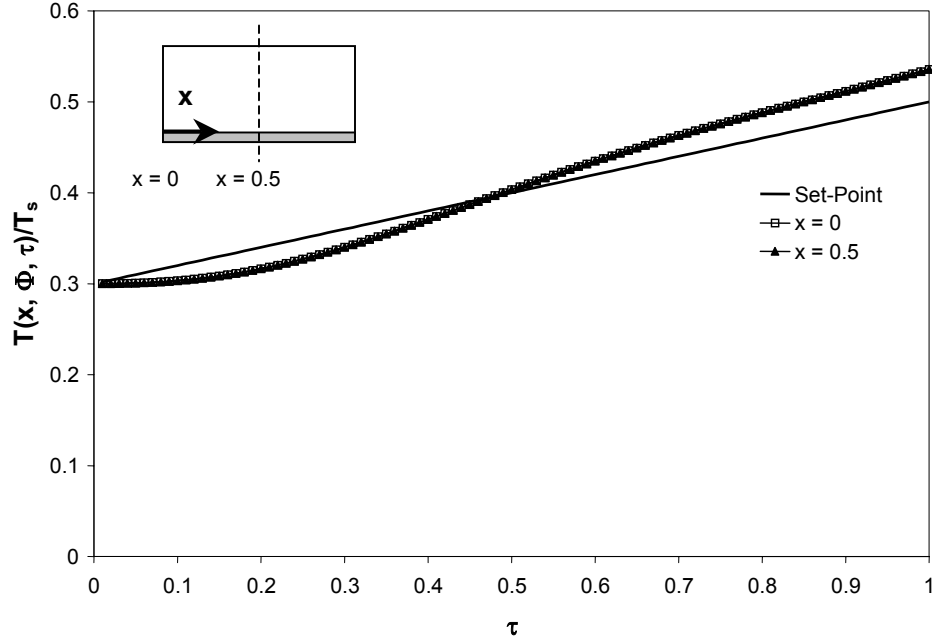


Figure 5.7: Variation of the design surface temperature from the set-point temperature throughout the process, for the problem shown in Fig. 5.5.

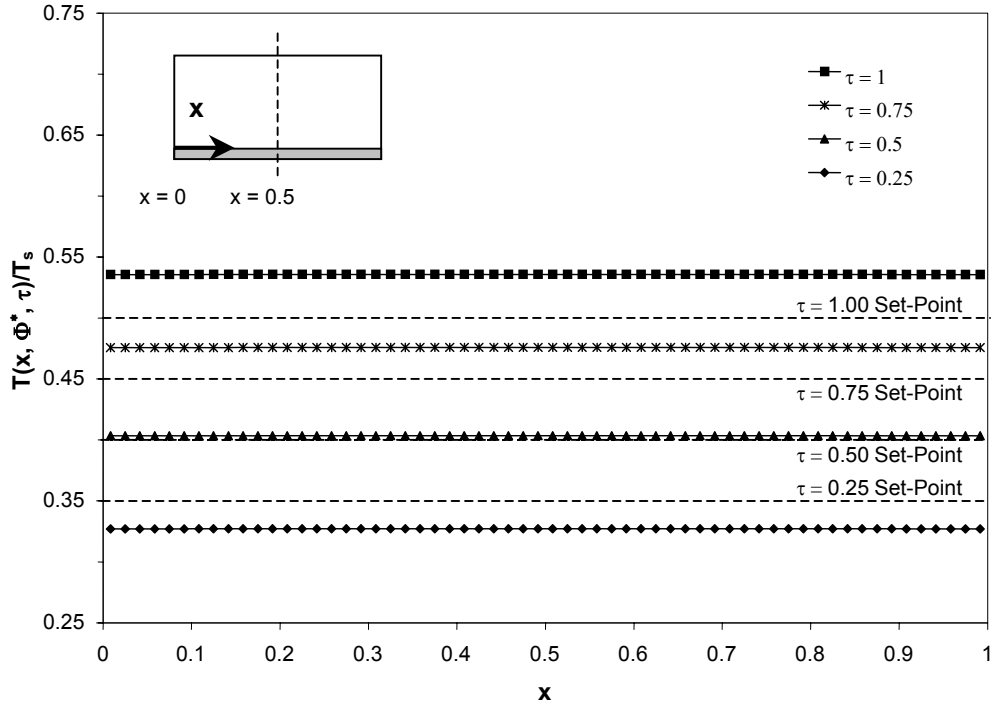


Figure 5.8: Temperature distribution over the design surface at different process times for the problem shown in Fig. 5.5.

Due to the approximations made in Eqs. (5.26) – (5.29), the transient temperature distribution found using the approach described above only approximates the analytical solution. In order to verify the grid-independence of the optimal solution, a grid refinement study was performed on  $F(\Phi^*)$ , as shown in Fig. 5.9. A total of  $N_t = 500$  time steps was used at each level of spatial refinement, and the grid-independent solution was taken to be the value of  $F(\Phi^*)$  calculated at the highest level of grid refinement,  $F_\infty(\Phi^*) \approx F_{N=480}(\Phi^*) = 2.283 \times 10^{-3}$ . Using the error defined in Eq. (2.35), the grid refinement error at  $N = 240$  is estimated to be 0.05%. Grid independence is also

demonstrated by examining the degree of energy conservation at different levels of refinement, defined as

$$\%EI(\Phi, N, N_t) = \frac{Q_{in}(\Phi) + Q_{conv}(\Phi, N, N_t) - Q_{stored}(\Phi, N, N_t)}{|Q_{in}(\Phi)| + |Q_{conv}(\Phi, N, N_t)| + |Q_{stored}(\Phi, N, N_t)|} \times 100 \%, \quad (5.41)$$

where  $Q_{in}(\Phi)$  is the energy added to the system by the heaters,  $Q_{stored}(\Phi, N, N_t)$  is the sensible energy stored by all the wall elements, and  $Q_{conv}(\Phi, N, N_t)$  is the net heat transfer from the fluid to the enclosure wall elements throughout the process.

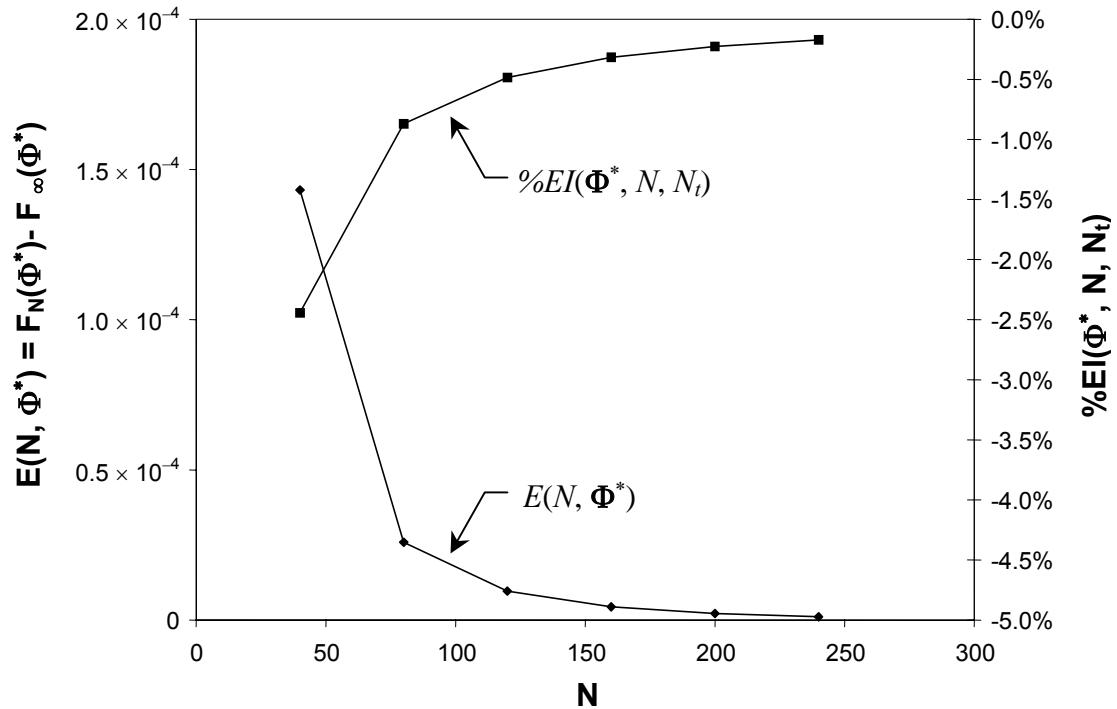


Figure 5.9: Spatial grid refinement study at  $\Phi^*$ , for the problem shown in Fig. 5.4, using  $N_t = 500$  time steps.

The solution is also sensitive to the number of time steps used to calculate the temperature distribution, since both the approximation of the time derivatives in Eq

(5.25) and the linearization scheme used to solve Eq. (5.32) become accurate only as  $\Delta t$  becomes small. In order to ensure that  $N_t$  was sufficiently large, a refinement study is performed on the temperature at the center and edge of the design surface obtained using the optimal transient heater settings at  $\tau = 0.5$ , which is shown in Fig. 5.10. A total of  $N = 240$  surface elements were used at each level of refinement to calculate the temperature distribution, and the grid independent temperatures were taken to be those calculated at the highest level of refinement,  $N_t = 1000$  time steps. Based on this refinement study, the temperatures calculated using  $N_t = 500$  time steps are within 0.1% of their grid-independent values.

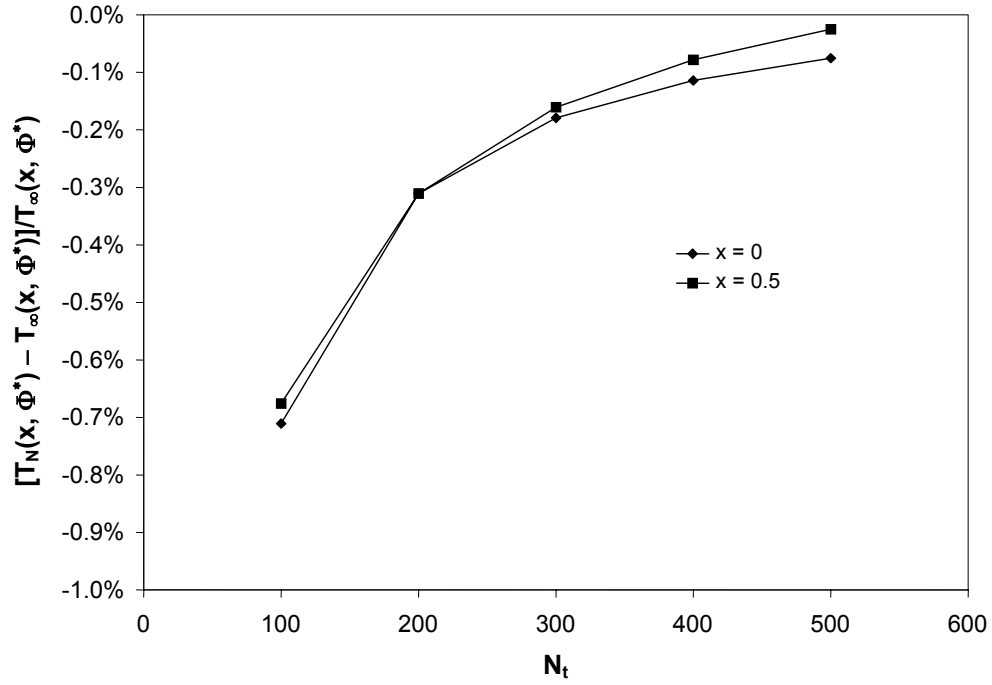


Figure 5.10: Temporal grid refinement study at  $\Phi^*$ , for the problem shown in Fig. 5.4, using  $N = 240$  wall elements.

### 5.3 Optimization Methodology for Low-Thermal Inertia Radiant Enclosures

While the optimization methodology presented in the previous section can be used to solve a wide range of radiant enclosure design problems involving transient and multimode heat transfer, in many cases it is difficult to implement because of the nonlinear nature of the governing equations. In particular, choosing an appropriate linearization scheme is not always intuitive, and the designer often resorts to selecting one by trial-and-error.

This section presents an optimization methodology that can be used to solve problems where thermal radiation is the only mode of heat transfer, and where the design surface is flat and is the only enclosure surface that stores sensible energy. In these cases, it is possible to rearrange the problem so that the governing equations are linear by specifying the temperature and the rate of temperature change over the design surface at each time step, in place of the adiabatic boundary condition. The heater settings that provide the required sensible energy by uniformly heating the design surface are then found using an optimization procedure similar to the one described in Chapter 2. Since this approach does not require a linearization scheme, it is much easier to implement than the optimization methodology presented in the previous section. Furthermore, because the heat flux distribution at each time step is found independently of the other time steps, there is no need to define a single function of  $\Phi$  that specifies the transient heater response throughout the process, like Eq. (5.37). Accordingly, this methodology is particularly well suited for solving problems where a fast heater response is required in order to produce rapid temperature changes in the design surface.



Although the problem is linearized by specifying the temperature boundary condition over the design surface, in the majority of applications the adiabatic boundary condition is enforced by insulating the underside of the design surface. Because of this, the quality of the optimal solution must be verified by first rewriting the problem so that the adiabatic boundary condition is specified over the design surface, and then solving the resulting nonlinear governing equations for the transient temperature distribution produced by the optimal heater settings. This design methodology therefore consists of two parts: the optimal heater settings are solved for in the first part of the procedure, and the resulting temperature distribution over the design surface is verified in the second part.

In this section, the optimization strategy is first presented, followed by a brief description of how the net radiant heat flux entering the design surface and the corresponding sensitivities are calculated. After that, the methodology used to calculate the transient temperature distribution as part of the validation procedure is presented. Finally, the design methodology is demonstrated by using it to find the heater settings that heat a silicon wafer uniformly according to a desired heating curve in an axisymmetric RTP furnace.

### 5.3.1 Optimization Strategy

As noted in the previous section, this class of problem has two design objectives: (a) to heat the design surface according to a prescribed heating curve, and (b) to maintain a uniform temperature over the design surface throughout the process. The first objective could be satisfied at the  $t^{th}$  time step by minimizing the objective function similar to the one defined in Eq. (5.1),

$$F_1[\Phi(t_t), t_t] = \{Q_{t-1 \rightarrow t}^{added}[\Phi(t_t), t_t] - Q_{t-1 \rightarrow t}^{target}(t_t)\}^2. \quad (5.42)$$

Unlike Eq. (5.2), however, the term  $Q_{t-1 \rightarrow t}^{added}[\Phi(t_t), t_t]$  is found by

$$Q_{t-1 \rightarrow t}^{added}[\Phi(t_t), t_t] = \sum_{i=1}^N q_{rad,i}[\Phi(t_t), t_t] \Delta A_i \Delta t_t, \quad (5.43)$$

where  $q_{rad,i}[\Phi(t_t), t_t]$  is the net radiant heat flux entering the  $i^{th}$  element over the  $t^{th}$  time step (which is equal to the difference between incoming and outgoing radiosity), and  $Q_{t-1 \rightarrow t}^{target}(t_t)$  is defined in Eq. (5.3). Note that in this methodology,  $\Phi$  is a function of  $t_t$  since a different set of design parameters is used to control the heat flux distribution over the heater surface at each time step. Minimizing  $F_1[\Phi(t_t), t_t]$  at each time step would result in an average design surface temperature that closely follows the set-point temperature throughout the process.

Assuming that the design surface properties are uniform, and in the absence of conduction and convection heat transfer, a uniform temperature increase over the design surface at each time step is realized by ensuring that a uniform amount of incident radiation is absorbed by each design surface element. This condition is satisfied by minimizing an objective function analogous to Eq. (5.5),

$$F_2[\Phi(t_t), t_t] = \frac{1}{N_{DS}} \sum_{i=1}^{N_{DS}} \{q_{rad,i}[\Phi(t_t), t_t] - \bar{q}_{rad}[\Phi(t_t), t_t]\}^2, \quad (5.44)$$

as shown in Fig. 5.11, where  $\bar{q}_{rad}[\Phi(t_i), t_i]$  is the average radiant heat flux absorbed by the design surface elements over the  $t^{th}$  time step.

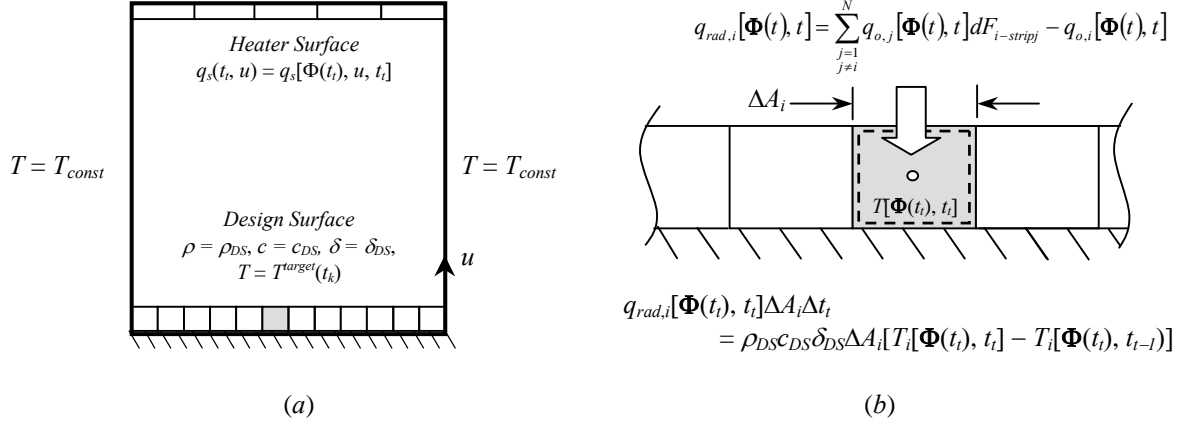


Figure 5.11: Example of a low-thermal inertia radiant enclosure: (a) enclosure boundary conditions and (b) relationship between radiosity, net radiant heat flux, and temperature increase.

An acceptable compromise between the competing design objectives is again found by minimizing a hybrid objective defined using the above two objective functions,

$$F[\Phi(t_i), t_i] = C \{ \gamma F_1[\Phi(t_i), t_i] + (1 - \gamma) F_2[\Phi(t_i), t_i] \}, \quad (5.45)$$

where  $\gamma$  is chosen according to the relative importance of the design objectives, and  $C$  is used to improve the objective function scaling. This objective function is minimized at each time step using the quasi-Newton method, which requires evaluations of  $F[\Phi(t_i), t_i]$  and  $\nabla F[\Phi(t_i), t_i]$ . The terms in the gradient vector are found from

$$\frac{\partial F[\Phi(t_t), t_t]}{\partial \Phi_p} = C \left\{ \gamma \frac{\partial F_1[\Phi(t_t), t_t]}{\partial \Phi_p} + (1-\gamma) \frac{\partial F_2[\Phi(t_t), t_t]}{\partial \Phi_p} \right\}, \quad (5.46)$$

where

$$\begin{aligned} \frac{\partial F_1[\Phi(t_t), t_t]}{\partial \Phi_p} = & 2 \left[ \sum_{i=1}^N q_{rad,i}[\Phi(t_t), t_t] \Delta A_i \Delta t_k - Q_{t-1 \rightarrow t}^{target}(t_t) \right] \\ & \times \left[ \sum_{i=1}^N \frac{\partial q_{rad,i}[\Phi(t_t), t_t]}{\partial \Phi_p} \Delta A_i \Delta t_t \right], \end{aligned} \quad (5.47)$$

and

$$\begin{aligned} \frac{\partial F_2[\Phi(t_k), t_t]}{\partial \Phi_p} = & \frac{2}{N_{DS}} \sum_{i=1}^N \{ q_{rad,i}[\Phi(t_t), t_t] - \bar{q}_{rad}[\Phi(t_t), t_t] \} \\ & \times \left\{ \frac{\partial q_{rad,i}[\Phi(t_t), t_t]}{\partial \Phi_p} - \frac{\partial \bar{q}_{rad}[\Phi(t_t), t_t]}{\partial \Phi_p} \right\}. \end{aligned} \quad (5.48)$$

A method for calculating  $q_{rad,i}[\Phi(t_t), t_t]$  and  $\partial q_{rad,i}[\Phi(t_t), t_t]/\partial \Phi_p$  is presented in the next section.

### 5.3.2 Calculation of the Net Radiant Heat Flux Distribution and Sensitivities

The radiant heat flux entering each wall element,  $q_{rad,i}[\Phi(t_t), t_t]$ , and the corresponding sensitivities,  $\partial q_{rad,i}[\Phi(t_t), t_t]/\partial \Phi_p$ , are found from the radiosity distribution and the radiosity sensitivity distributions through a procedure similar to the one described in Chapter 2.

As before, the first step is to identify a suitable parametric representation for the enclosure geometry, thermodynamic boundary conditions, and surface properties. In particular, the heat flux distribution over the heater surfaces is specified as a function of  $t_t$  and  $\Phi(t_t)$ , while the temperature over the design surface is set equal to the set-point temperature,  $T[u, \Phi(t_t), t_t] = T^{target}(t_t)$ .

Once the enclosure is represented parametrically, the integral equation governing the radiosity distribution can then be written as

$$q_o[u, \Phi(t_i), t_i] = b[u, \Phi(t_i), t_i] + g(u) \int_a^b q_o[u', \Phi(t_i), t_i] k(u, u') du', \quad (5.49)$$

with  $b[u, \Phi(t_i), t_i]$  and  $g(u)$  defined in Eqs. (2.3) and (2.4), respectively. By differentiating Eq. (5.49) with respect to  $\Phi_p$ , it can be shown that the radiosity sensitivities are governed by a related integral equation,

$$\frac{\partial q_o[u, \Phi(t_i), t_i]}{\partial \Phi_p} = \frac{\partial b[u, \Phi(t_i), t_i]}{\partial \Phi_p} + g(u) \int_a^b \frac{\partial q_o[u', \Phi(t_i), t_i]}{\partial \Phi_p} k(u, u') du'. \quad (5.50)$$

Finally, assuming that the underside of the design surface is adiabatic, the net radiant heat flux entering an infinitesimal wall element is given by the difference between the incident radiosity from the other locations on the enclosure surface and the radiosity leaving the infinitesimal surface element,

$$q_{rad}[u, \Phi(t_i), t_i] = \int_a^b q_o[u', \Phi(t_i), t_i] k(u, u') du' - q_o[u, \Phi(t_i), t_i], \quad (5.51)$$

while the corresponding sensitivities are found by

$$\frac{\partial q_{rad}[u, \Phi(t_i), t_i]}{\partial \Phi_p} = \int_a^b \frac{\partial q_o[u', \Phi(t_i), t_i]}{\partial \Phi_p} k(u, u') du' - \frac{\partial q_o[u, \Phi(t_i), t_i]}{\partial \Phi_p}. \quad (5.52)$$

As previously noted, analytical solutions to these types of integral equations are not usually tractable, so the radiosity and the radiosity sensitivity distributions are estimated numerically. By first discretizing the parametric region into  $N$  elements and then following the steps described in Eqs. (2.5) and (2.6), Eq. (5.49) is transformed into a matrix equation that can be solved for the radiosity values at discrete locations on the enclosure surface,

$$\mathbf{A}b[\Phi(t_i), t_i] = \mathbf{x}[\Phi(t_i), t_i], \quad (5.53)$$

where  $b_i[\Phi(t_i), t_i] = b[u_i, \Phi(t_i), t_i]$ ,  $x_i[\Phi(t_i), t_i] = q_o[u_i, \Phi(t_i), t_i]$ , and  $\mathbf{A}$  is defined by Eq. (2.8). Using a similar procedure, it can be shown that the radiosity sensitivities are found by solving

$$\mathbf{A}b'[\Phi(t_i), t_i] = \mathbf{x}'[\Phi(t_i), t_i], \quad (5.54)$$

where  $b'_i[\Phi(t_i), t_i] = \partial b[u_i, \Phi(t_i), t_i]/\partial \Phi_p$ , and  $x'_i[\Phi(t_i), t_i] = \partial q_o[u_i, \Phi(t_i), t_i]/\partial \Phi_p$ .

The radiosity values are then used to calculate the net radiative heat flux entering a wall element,  $q_{rad,i}[\Phi(t_i), t_i]$ ,

$$q_{rad,i}[\Phi(t_i), t_i] = \sum_{\substack{j=1 \\ j \neq i}}^N q_{oj}[\Phi(t_i), t_i] dF_{i-stripj} - q_{oi}[\Phi(t_i), t_i], \quad (5.55)$$

and the corresponding sensitivities,

$$\frac{\partial q_{rad,i}[\Phi(t_i), t_i]}{\partial \Phi_p} = \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\partial q_{oj}[\Phi(t_i), t_i]}{\partial \Phi_p} dF_{i-stripj} - \frac{\partial q_{oi}[\Phi(t_i), t_i]}{\partial \Phi_p}. \quad (5.56)$$

Finally, the solutions for  $q_{rad}[\Phi(t_i), t_i]$  and  $\partial q_{rad}[\Phi(t_i), t_i]/\partial \Phi_p$  are substituted into Eqs. (5.42)-(5.48) to calculate  $F[\Phi(t_i), t_i]$  and  $\partial F[\Phi(t_i), t_i]/\partial \Phi_p$ .

### 5.3.3 Verification of Optimal Heater Settings

The methodology presented in the previous section assumes that the temperature boundary condition is enforced over the design surface throughout the process, and a near-uniform incident radiant heat flux distribution over the design surface that supplies the required sensible energy is found by minimizing the objective function defined by Eq. (5.45). Although the minimum value of the objective function at any given time step will

be very small, it is unlikely that  $F_1[\Phi^*(t_i), t_i]$  or  $F_2[\Phi^*(t_i), t_i]$  will exactly equal zero. This means that either too much or too little sensible energy has been added to the design surface at  $t_i$ , and that the incident radiant heat flux distribution is not precisely uniform.

Thus, if these optimal heater settings were used to heat an insulated design surface, the average design surface temperature would deviate slightly from the set-point temperature and the temperature distribution over the design surface would also be slightly non-uniform at any given instant. It is therefore often necessary to calculate the transient temperature distribution over an adiabatic design surface in order to verify the quality of the optimal solution.

At any time  $t_i$  during the process, it can be shown that the radiosity distribution over the enclosure surfaces is governed by modified versions of Eqs. (2.2) – (2.4),

$$q_o[u, \Phi^*(t_i), t_i] = b[u, \Phi^*(t_i), t_i] + g(u) \int_a^b q_o[u', \Phi^*(t_i), t_i] k(u, u') du', \quad (5.57)$$

where  $b[u, \Phi^*(t_i), t_i] = q_s[u, \Phi^*(t_i), t_i]$  and  $g(u) = 1$  if  $u$  lies on the heater surface, and either

$$b[u, \Phi^*(t_i), t_i] = \varepsilon(u) \sigma T^4[u, \Phi^*(t_i), t_i], \quad g(u) = 1 - \varepsilon(u), \quad (5.58)$$

or equivalently

$$b[u, \Phi^*(t_i), t_i] = -\rho(u) c(u) \delta(u) \frac{\partial T[u, \Phi^*(t_i), t_i]}{\partial t}, \quad g(u) = 1, \quad (5.59)$$

if  $u$  lies on the design surface.

If the adiabatic boundary condition is specified over the design surface, however, then both  $T[u, \Phi^*(t_i), t_i]$  and  $\partial T[u, \Phi^*(t_i), t_i]/\partial t$  are unknown. Instead, the problem is solved at each time step starting at  $t = 1$  by first substituting  $T[u, \Phi(t_i), t_i] \approx T[u, \Phi(t_{i-1}), t_{i-1}]$  into Eq. (5.58) and then solving Eq. (5.57) numerically using the

procedure described in Chapter 2. Once this has been done, the resulting radiosity distribution is used to calculate the average radiant heat flux entering a design surface wall element during the  $t^{th}$  time step using Eq. (5.55), which in turn is used to find the corresponding temperature increase in each element of the design surface by

$$T_i[\Phi(t_t), t_t] = T_i[\Phi(t_{t-1}), t_{t-1}] + \frac{q_{rad,i}[\Phi(t_t), t_t] \Delta A_i \Delta t_t}{\rho_i c_i \delta_i}. \quad (5.60)$$

This new temperature is substituted into Eq. (5.58), and the procedure is repeated to find the temperature distribution at the next time step.

#### 5.3.4 Implementation and Demonstration

This design methodology is demonstrated by using it to calculate the heater settings for the RTP furnace shown in Fig. 5.12, which is based on design problem presented by Choi and Do (2001). The furnace is circular-cylindrical, with a height of  $H = 0.5 \text{ m}$  and a radius of  $R = 0.5 \text{ m}$ . The heating surface is at the top of the enclosure, and is composed of three concentric heaters that have negligible thermal inertia and an emissivity  $\varepsilon_{HS} = 0.9$ . The heat flux distribution over each of the heaters is assumed to be uniform and is controlled by one of the three design parameters at each time step. The design surface is a  $0.75 \text{ m}$  diameter  $\text{SiO}_2$  semiconductor wafer located on the bottom of the enclosure with properties summarized in Table 5.2. The remainder of the enclosure surface has an emissivity of  $\varepsilon_v = 0.4$  and is maintained at a uniform temperature of  $T_{wall} = 300 \text{ K}$  throughout the process using an active cooling system.

The objective of the design problem is to determine the transient heater settings that heat the semiconductor wafer according to a desired heating history, while



simultaneously maintaining a uniform temperature distribution throughout the process. The process consists of three stages: in the first stage, the silicon wafer is heated from an initial temperature of  $500\text{ K}$  to a temperature of  $1000\text{ K}$  over a 5 second time span. The wafer is maintained at this temperature for a further 5 seconds, and then cooled at a constant ramp rate from  $1000\text{ K}$  to  $850\text{ K}$  over 10 seconds.

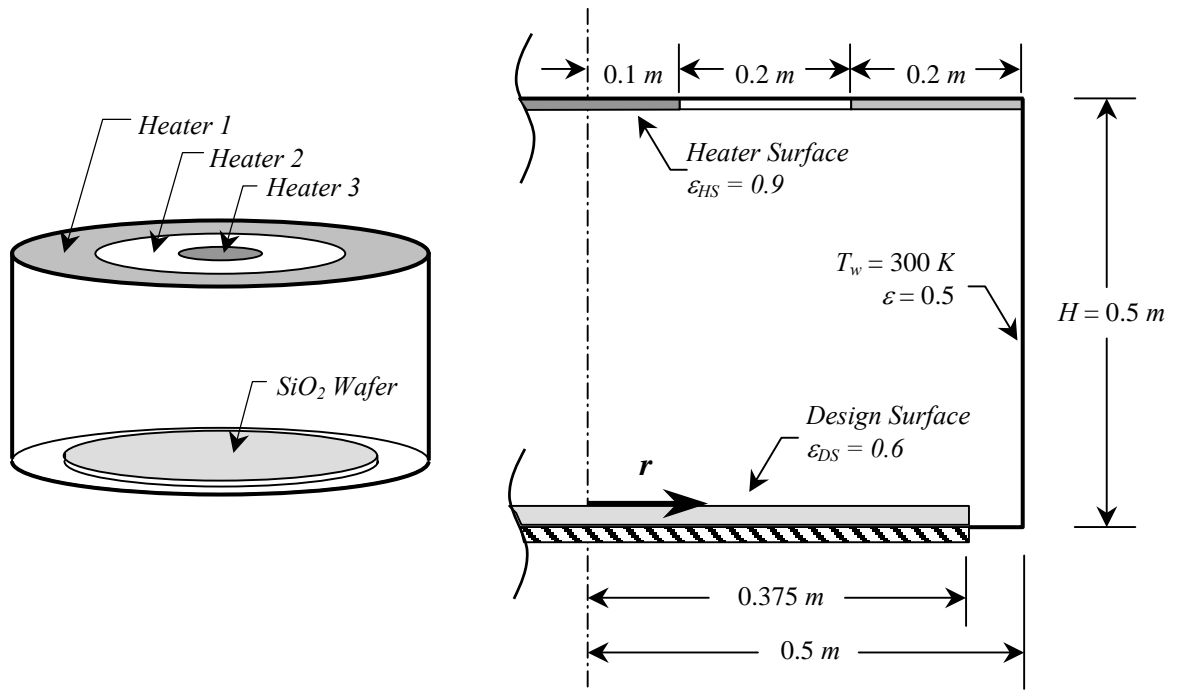


Figure 5.12: Rapid thermal processing furnace design problem.

$\rho$ [ $\text{kg/m}^3$ ]	$c$ [ $\text{J/kg K}$ ]	$\delta$ [ $\text{mm}$ ]	$\varepsilon$
2220	745	0.675	0.6

Table 5.2: Design surface properties for the problem shown in Fig. 5.15.

In this problem, temperature, heat flux, and time are all represented non-dimensionally by  $\hat{T}_i[\Phi(t_i), t_i] = T_i[\Phi(t_i), t_i]/T_s$ ,  $\hat{q}_{si}[\Phi(t_i), t_i] = q_{si}[\Phi(t_i), t_i]/\sigma T_s^4$ , and  $\tau = t/t_{max}$  respectively, with  $T_s = 1000\text{ K}$  and  $t_{max} = 20\text{ s}$ . The heat flux distribution over the heater surface at any non-dimensional time  $\tau_i$  is given by  $q_s[u, \Phi(t_i), t_i]/\sigma T_s^4 = \Phi_h(\tau_i)$ , where  $u$  lies on the  $h^{th}$  heater. The heater output range was enforced by constraining the design parameters to lie between  $0 \leq \Phi_h \leq 20$ ,  $h = 1 \dots 3$ , using an active set method.

The optimal heater settings were solved by minimizing the objective function defined in Eq. (5.45). Minimization was carried out at each time step using the quasi-Newton (BFGS) minimization technique. A total of  $N = 500$  surface elements were used to solve for the radiosity distribution; 400 of these were uniformly spaced annular elements on the upper and lower enclosure surfaces and the remaining 100 were ring elements on the side walls. Since the equations that govern the radiant heat flux entering the design surface during the optimization procedure are linear with respect to time, the solution accuracy is not affected by the number of time steps used to carry out the optimization; rather, it is necessary only to use a sufficient number of time steps in order to obtain a quasi-continuous heater response. On the other hand, the temperature distribution calculated during the validation procedure is sensitive to  $N_t$  since each element is assumed to have a constant temperature for the duration of each time step; thus, as  $N_t$  becomes larger, the error induced by this assumption becomes smaller. Accordingly,  $N_t = 100$  time steps are used to optimize the heater settings, and  $N_t = 2000$  time steps are used throughout the validation process.

In each case, minimization was started with  $\Phi^0(\tau_k) = \{1, 1, 1\}^T$ , and stopped when  $\nabla F_{FR}[\Phi^k(\tau_i)] < 10^{-6}$ . After several attempts, a good solution was found using  $C = 10^4$  and  $\gamma = 0.9999$ . (This value of  $\gamma$  is much smaller than the one used to optimize the previous problem because  $Q_{t-1 \rightarrow t}^{added}[\Phi(t_i), t_i] - Q_{t-1 \rightarrow t}^{target}(t_i)$  is evaluated over only the current time step, while in the previous methodology it is integrated over the entire process.) The optimal heater settings are shown in Fig. 5.13, and the corresponding values of  $F(\Phi^*)$  and the number of iterations required to reach  $\Phi^*$  at each time step are shown in Fig. 5.14. As shown in Fig. 5.13, during the initial heating phase from  $\tau = 0$  to  $\tau = 0.25$ , the outer annular heater has the highest output (since a substantial amount of the radiant heat produced by the outer heater is lost to the cold side wall) while the middle heater is turned off and the inner heater is on at a low setting.

The output from the outer heater then drops while that of the inner heater increases as the wafer is held at a constant temperature from  $\tau = 0.25$  to  $\tau = 0.5$ . Throughout the cooling phase, from  $\tau = 0.5$  to  $\tau = 1$ , the wafer is radiantly cooled by the side walls and the outer annular heater is deactivated. In the first part of the cooling phase, from  $\tau = 0.5$  to  $\tau = 0.77$ , the cooling rate is moderated by the action of the inner heater, and the middle heater is briefly activated in order to maintain a uniform temperature distribution over the heater. Near the end of the cooling process, from  $\tau = 0.77$  to  $\tau = 1$ , all three heaters are deactivated and the cooling rate and temperature distribution over the wafer is driven entirely by the radiant exchange between the wafer and the side wall.

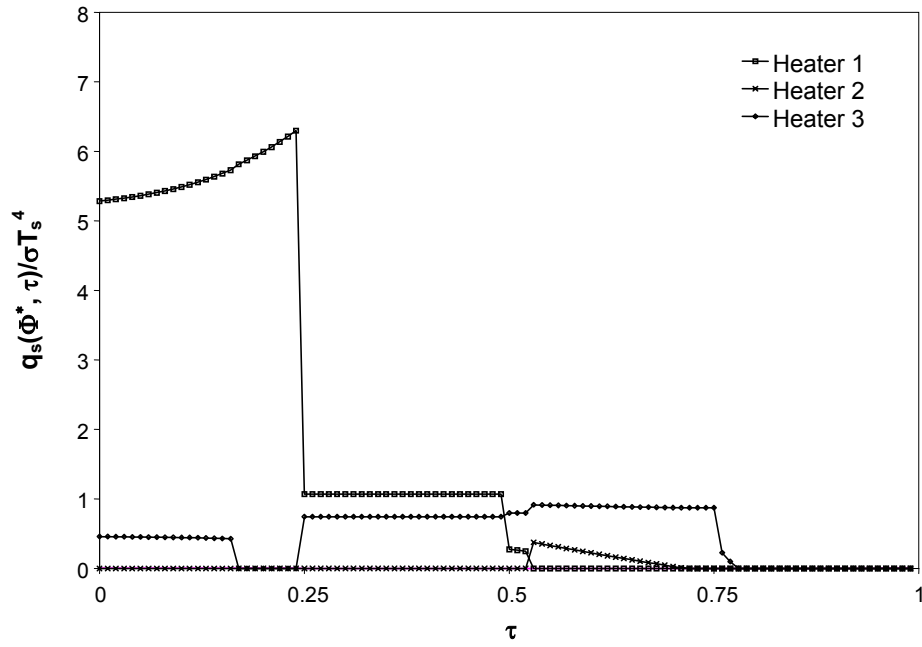


Figure 5.13: Optimal heater settings obtained using  $N = 500$  wall elements, and  $N_t = 100$  time steps, for the problem shown in Fig. 5.12.

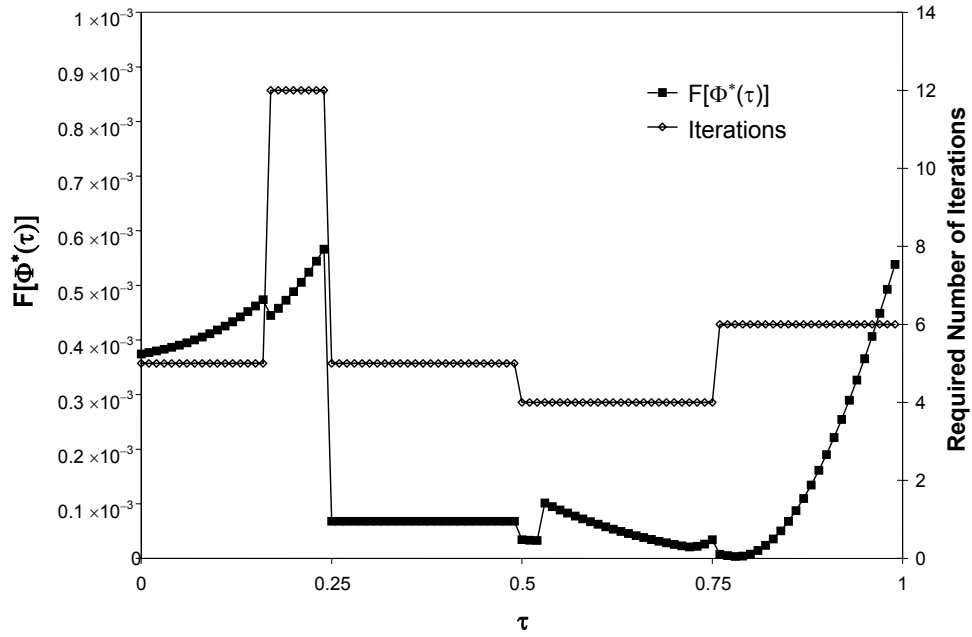


Figure 5.14: Values of  $F[\Phi^*(\tau)]$  corresponding to the heater settings shown in Fig. 5.13, and the number of time steps required to reach  $\Phi^*(\tau)$ .

The transient temperature distribution over the semiconductor wafer is shown in Figs. 5.15 and 5.16. Figure 5.15 shows that the maximum and minimum wafer temperatures are at the center and edge, respectively, closely match the set-point temperature throughout the process. The maximum deviation between the average wafer temperature and the set-point temperature is 1.09%, occurring at the end of the process. As shown in Fig. 5.16, a near-uniform temperature distribution is maintained over the wafer at all times. The maximum standard deviation of the wafer temperatures is 0.63%, which occurs at  $\tau = 0.25$ .

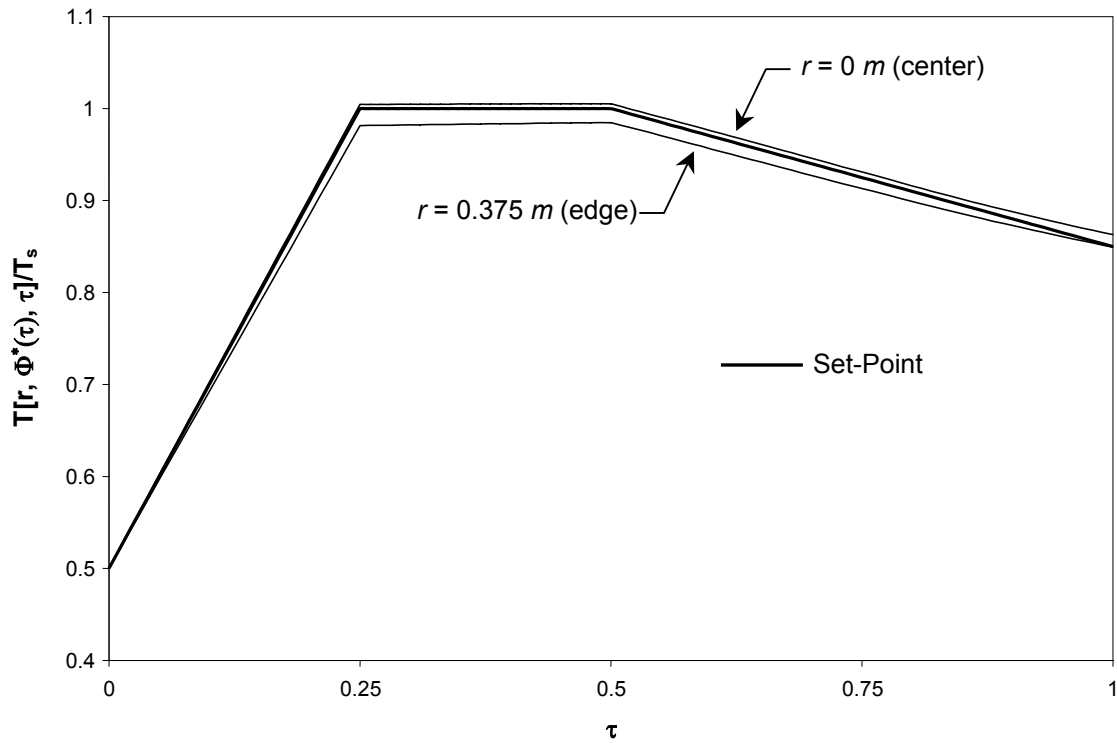


Figure 5.15: Variation of the design surface temperature from the set-point temperature throughout the process, for the problem shown in Fig. 5.12.

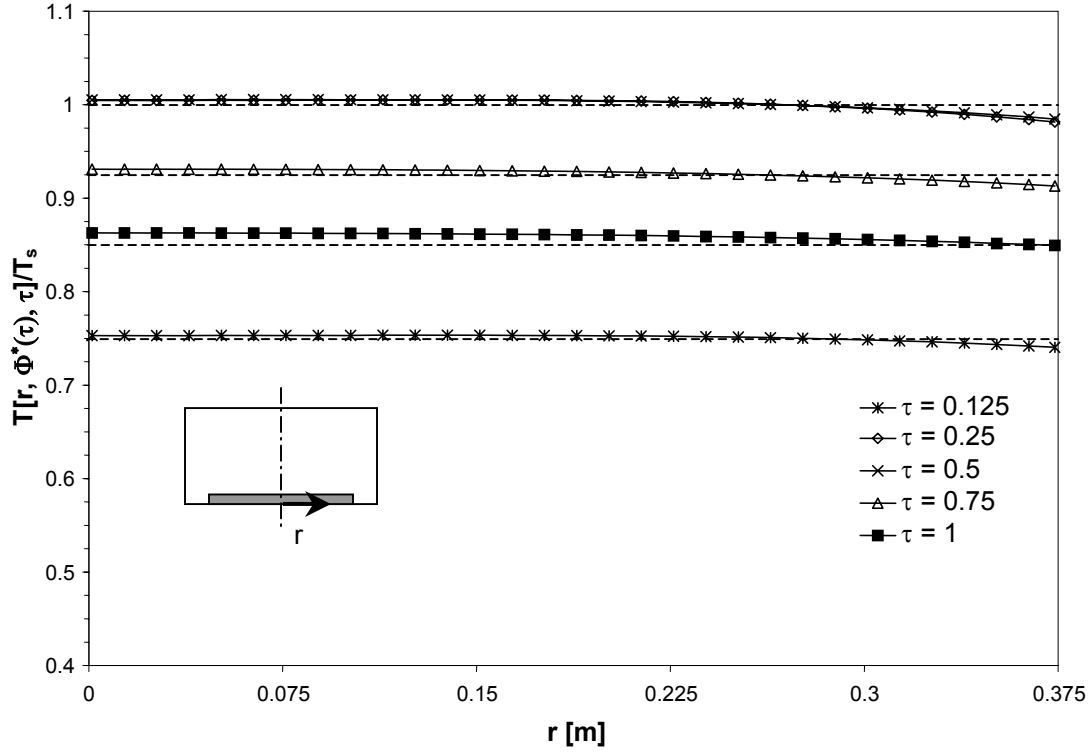


Figure 5.16: Non-dimensional temperature distribution over the design surface at different process times, for the problem shown in Fig. 5.12.

As previously noted, assuming a uniform radiosity distribution over each surface element induces an error in the calculation of the heat flux distribution over the design surface; this error becomes small as more elements are used in the analysis and approaches zero as  $N$  becomes large. In order to verify that a sufficient number of surface elements was used in the optimization procedure, a refinement study is performed on the average value of  $F[\Phi^*(\tau_i), \tau_i]$  obtained at all time steps,  $F_{ave}(\Phi^*)$ , and is shown in Fig. 5.17. (The term  $\Phi^*$  is a matrix containing the optimal design parameters from each time step.) The grid independent value of  $F_{ave}(\Phi^*)$  was taken to be that obtained at the

highest level of refinement,  $F_{ave,\infty}(\Phi^*) = F_{ave,N=2000}(\Phi^*)$  and the refinement error was then defined as

$$E_{Fave}(N, N_t, \Phi^*) = F_{ave,N}(\Phi^*) - F_{ave,\infty}(\Phi^*). \quad (5.61)$$

Using this analysis, the refinement error in  $F_{ave,N=500}(\Phi^*)$  calculated with  $N_t = 100$  time steps was estimated to be less than 0.2 %.

In order to determine the error induced by using a finite number of surface elements throughout the verification procedure, a second grid refinement study was performed on the degree of energy conservation, defined as

$$\%EI(\Phi^*, N) = \frac{Q_{in,N}(\Phi^*) - Q_{out,N}(\Phi^*) + Q_{stored,N}(\Phi^*)}{|Q_{in,N}(\Phi^*)| + |Q_{out,N}(\Phi^*)| + |Q_{stored,N}(\Phi^*)|} \times 100\%, \quad (5.62)$$

where  $Q_{in,N}(\Phi^*)$  is the total energy added by the heaters,  $Q_{out,N}(\Phi^*)$  is the total energy removed by the isothermal walls, and  $Q_{stored,N}(\Phi^*)$  is the sensible energy stored by the design surface throughout the process. Figure 5.17 shows that a sufficient level of energy conservation is achieved using  $N=500$  elements.

A second error is induced into the verification procedure by assuming that each element has a constant temperature over each time step; again, this error tends to zero as  $N_t$  becomes large. In order to verify that a sufficient number of time steps was used in the verification procedure, a refinement study was performed on the non-dimensional temperatures at the center ( $r = 0$  m) and outer edge of the wafer ( $r = 0.75$  m) at  $\tau = 0.26$ , obtained using  $N = 500$  surface elements.

The grid-independent values are taken to be those obtained using the largest number of time steps,  $N_t = 5 \times 10^4$ . The refinement error in the temperature of the  $i^{th}$  element is then

$$E_{Ti}(N, N_t, \Phi^*, \tau) = \hat{T}_{i,N}(\Phi^*, \tau) - \hat{T}_{i,\infty}(\Phi^*, \tau). \quad (5.63)$$

As shown in Fig. 5.18, using  $N_t = 2000$  time steps results in an error of approximately 0.01%.

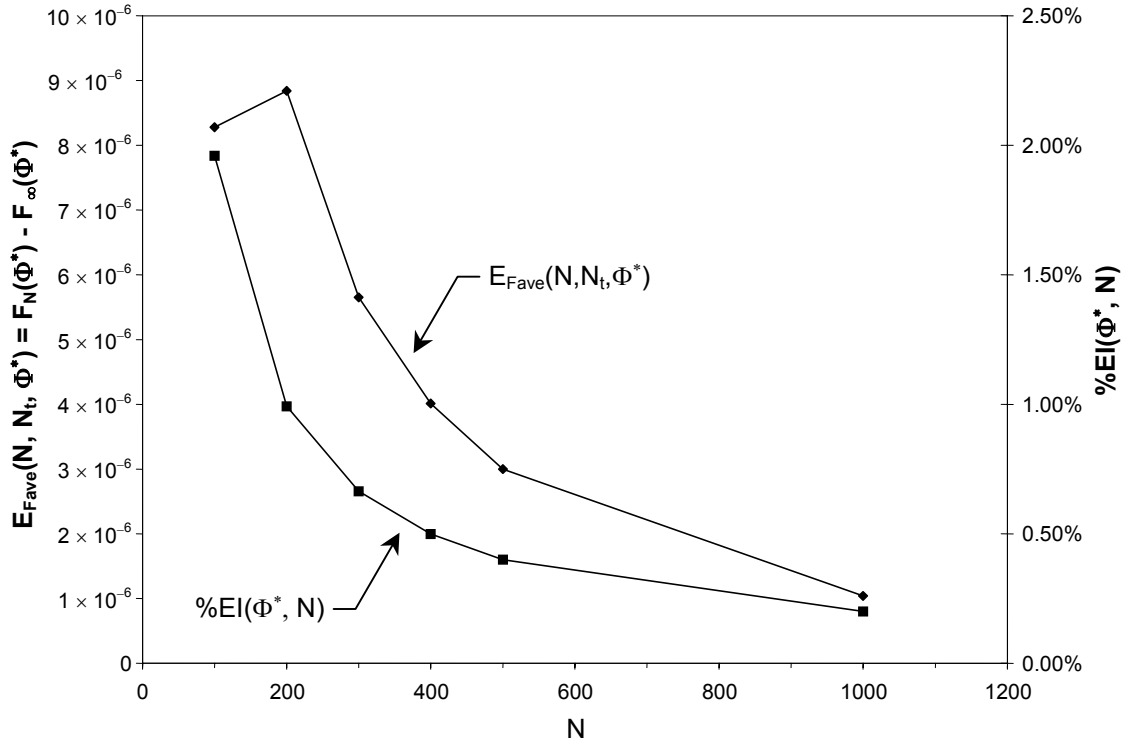


Figure 5.17: Refinement study on the number of surface elements used in the optimization and verification procedures.



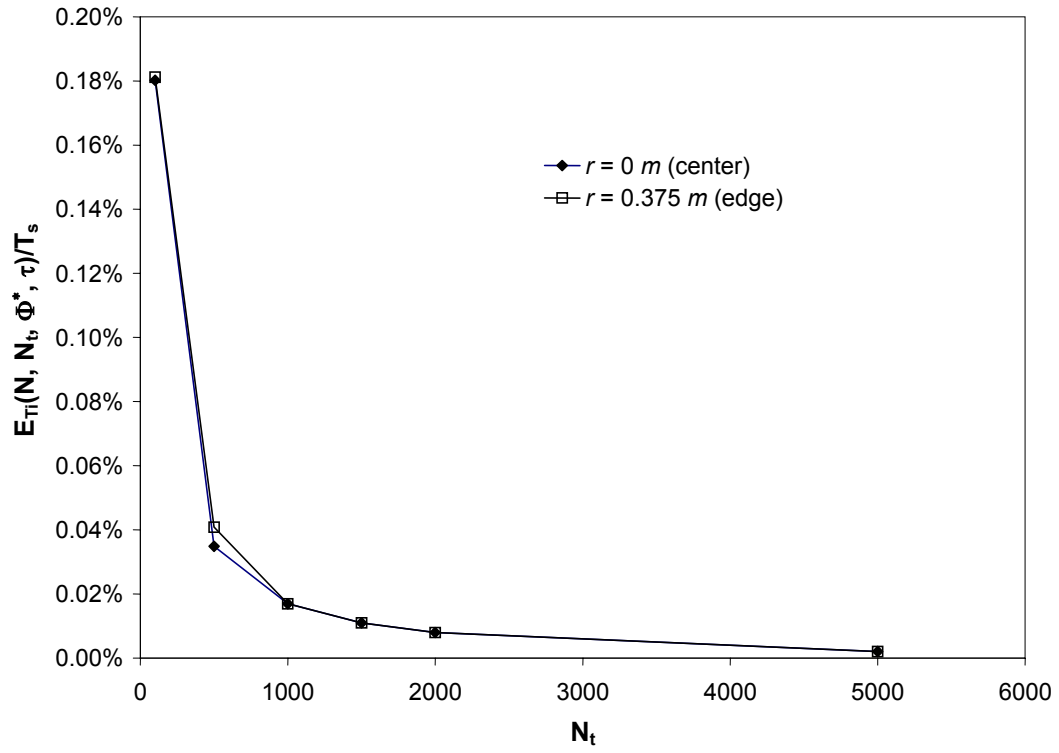


Figure 5.18: Refinement study on the number of time steps used in the verification procedure.

## 5.4 Conclusions

The objective of many radiant enclosure design problems is to find the transient heat flux distribution over the heater surface that heats a product according to a desired temperature history, while simultaneously maintaining a uniform temperature distribution at any given instant in order to ensure product homogeneity. An *a priori* estimate of this heat flux distribution is used to size the radiant heaters when designing these enclosures, and also enables the use of high-gain controllers under operating conditions.

This chapter presented two design methodologies that can be used to solve this type of design problem. Both methods work by minimizing a hybrid objective function composed of terms relating to the amount of sensible energy added to the design surface and the temperature variation within the product throughout the process.

The first methodology is generic in the sense that conduction heat transfer through the enclosure walls, convection heat transfer between the walls and the surrounding fluids, and sensible energy storage within the walls are all accounted for in the heat transfer model, and it can therefore be applied to solve many different problems. In this methodology, the objective function terms are integrated over the process duration and a single set of design parameters controls the transient heat flux distribution over the heater surface at all times. The objective function is defined in terms of the temperature and temperature sensitivities evaluated at discrete locations over the design surface at different process times, which in turn are found by solving the set of nonlinear governing equations numerically using a problem-specific linearization scheme. This design methodology was demonstrated by using it to solve for the optimal transient heater

settings in an annealing furnace. In this example, the transient heat output from each heater was governed by a Bézier spline having design parameters as its coefficients.

The second methodology is used to solve problems where thermal radiation is the only mode of heat transfer and where the design surface is flat and is the only part of the enclosure that stores sensible energy. First, the governing equations are transformed from a nonlinear to a linear form by specifying the temperature and rate of temperature change over the design surface. Once this is done, the optimal heater settings are then found by minimizing an objective function at every time step involving the net radiant heat flux entering the design surface. Since the adiabatic boundary condition is enforced over the design surface in most practical applications, however, the quality of the optimal heater settings must be verified by calculating the transient temperature distribution over the design surface that is produced using the optimal heater settings.

Although this design methodology cannot be used to solve as many problems as the generic approach, it is usually easier to implement because the analyst does not need to select a linearization scheme in order to solve the problem. Moreover, this methodology is more suitable for low-thermal inertia applications where a quick heater response time is needed to produce rapid changes in the design surface temperature.

## **Chapter 6:**

## **Conclusions**

## 6.1 Conclusions

Design problems involving radiant enclosures are routinely encountered in many different industrial applications. In most problems, radiant heaters on the heater surface are used to impose desired temperature and heat flux distributions over the product located on the design surface; usually, these distributions need to be uniform in order to ensure product homogeneity.

When designing the radiant enclosure, the designer's objective is to identify an enclosure geometry and heat flux distribution over the heater surface that produces the desired conditions over the design surface. Traditionally, this has been done using a forward "trial-and-error" design methodology, where the designer repeatedly adjusts the design relying solely on his or her experience until the design surface conditions match the desired conditions within a specified tolerance. Due to the complicated nature of radiant exchange between the enclosure surfaces, however, an intuitive understanding of how to modify the design in order to improve the solution is elusive even to the most experienced designer. Because of this, the forward design methodology usually requires many iterations and consequently a lot of design time to carry out, and the procedure is usually stopped as soon as the conditions produced over the design surface match the desired conditions within an allowable tolerance, well before the design has been optimized.

Design problems involving radiant enclosures containing surfaces with directionally dependent properties and those with transient and multimode heat transfer effects are particularly complex, and are thus ill-suited to the forward methodology.

Because of this, designers often use approximate analytical techniques based on non-imaging optics to solve the former type of design problem, and control algorithms to solve the latter. Although these strategies usually produce better solutions and require far less design time than the forward methodology, they are also quite complex and the designer must possess specialized skills in order to carry out the analyses. Furthermore, both of these methods are based on simplified heat transfer models, which restrict solution accuracy.

Designers have recently applied inverse design methodologies to solve radiant enclosure design problems. In this approach, both the desired temperature and heat flux distributions are explicitly enforced over the design surface, while the conditions over the heater surface are unspecified. When written in this form, the design problem is mathematically ill-posed because there may exist multiple solutions that produce the desired conditions over the design surface, or no solution at all. Discretizing the governing equations results in an ill-conditioned (usually singular) system of equations that generally has an unequal number of equations and unknowns. Various regularization techniques are then used to find a solution that is smooth and also produces a sufficiently small residual when substituted into the system of equations.

The inverse design methodology usually requires far less design time than the forward “trial-and-error” methodology, and in most cases the heat flux and temperature distributions over the design surface closely match the desired distributions. A major drawback of this approach, however, is that it is very difficult to enforce design constraints, and because of this solutions are often found that are difficult to implement in

an industrial setting. For example, solutions obtained using this technique often have regions of negative heat flux over the heater surface; in most cases, the designer sets the heat flux output over these regions equal to zero and incurs the resulting error over the design surface. It is also sometimes difficult to rearrange a radiant enclosure design problem into an explicit form that can be solved using a regularization technique, and for this reason inverse design methodologies have not been used to solve design problems where the enclosure geometry is variable.

The optimization design methodology overcomes many drawbacks of these other design methodologies. It is similar to the forward design methodology in the sense that it works by repeatedly analyzing and then modifying trial designs until an acceptable solution is identified; unlike the forward design methodology, however, the design modifications are carried out systematically. This is made possible by first rewriting the design problem as a minimization problem by specifying one of the desired thermodynamic boundary conditions (usually the temperature distribution) over the design surface and then using the other to define an objective function,  $F(\Phi)$ , which quantifies the “goodness” of the design configuration. In most cases,  $F(\Phi)$  is the variance between the heat flux distribution and the desired heat flux distribution, evaluated at discrete locations over the design surface. The variables in the design problem, most often the enclosure geometry and the heat flux distribution over the heater surface, are controlled by a set of design parameters contained in the vector  $\Phi$ .

The optimal design is then found by minimizing  $F(\Phi)$  through repeated iteration. At the  $k^{th}$  iteration, the heat flux distribution over the design surface is first solved and

then used to evaluate  $F(\Phi^k)$ . Next,  $\nabla F(\Phi)$  and sometimes  $\nabla^2 F(\Phi)$  are formed, either from the heat flux sensitivities or using a finite difference approximation if these sensitivities are unavailable. Finally, a new set of design parameters,  $\Phi^{k+1}$ , is calculated based on the objective function sensitivities at  $\Phi^k$ . This procedure stops once a local minimum has been reached; the set of design parameters  $\Phi^*$  that minimizes  $F(\Phi)$  corresponds to the design configuration that best satisfies the desired conditions over the design surface.

Because the design parameters are systematically updated based on the objective function curvature, the optimization design methodology finds a much better solution in less time than the forward design methodology, which relies solely on the designer's experience and intuition to modify the design at each iteration. Also, since the enclosure problem is solved in its "natural" implicit form, it is possible to accommodate complex heat transfer models. Because of this, many different types of enclosure problems can be solved using this approach, and there is no need for the simplifying assumptions required in analyses involving non-imaging optics and control algorithms. Finally, unlike the inverse design methodology, design constraints can be enforced throughout the optimization process; in radiant enclosure design problems, constraints are most often used to ensure that the heat flux distribution over the heater surface is nonnegative, to force the enclosure geometry lie within a set of maximum dimensions, and to maintain an unobstructed enclosure geometry throughout the analysis.

In this dissertation, five optimization methodologies were presented for solving different types of radiant enclosure design problems. Chapter 2 described an



optimization methodology for finding a heat flux distribution over the heater surface that produces the desired temperature and heat flux distributions over the design surface in diffuse-walled enclosures with fixed geometry. The heat flux distribution is solved with infinitesimal-area method and the heat flux sensitivities are found efficiently by post-processing. Optimization was carried out using Newton's method and design constraints were enforced through an active set method. This optimization methodology was demonstrated by using it to solve an example design problem, and the solution quality was shown to be superior to that obtained using an inverse (TSVD) design methodology.

This optimization methodology was then extended in Chapter 3 to solve diffuse-walled enclosure design problems where the enclosure geometry is a design variable. Although any type of parametric representation can be employed to represent the enclosure geometry, Bézier curves have several inherent advantages in the context of geometric optimization. The heat flux and heat flux sensitivities over the design surface are found using a method similar to the one described in the previous chapter, and the objective function is again minimized using Newton's method. The active set method enforces design constraints that limit the enclosure size and ensure that the enclosure remains unobstructed throughout the optimization process, which is necessary in order to ensure that the objective function is continuously differentiable with respect to the design parameters.

Chapter 4 presented an optimization methodology for designing the geometry and heater settings of radiant enclosures containing surfaces having directionally-dependent optical properties. In this methodology, the heat flux distribution over the design surface

is solved by performing a Monte Carlo analysis based on exchange factors, which induces a statistical uncertainty in the solution. Because of this, the objective function is minimized using the Kiefer Wolfowitz method, a gradient-based algorithm developed specifically to optimize stochastic systems where unbiased estimates of the gradient vector are unavailable.

Finally, two optimization methodologies were presented in Chapter 5 that can be used to solve for the heater settings in diffuse-walled radiant enclosures with transient heat transfer effects. The first method is “generic” in the sense that it takes into account conduction heat transfer through the enclosure walls, convection heat transfer between the walls and the surrounding fluids, and sensible energy storage in the enclosure surfaces, and it can therefore be used to solve many different design problems. The transient heater outputs are controlled throughout the process with a single set of design parameters, which is done by setting the transient heat output of a single heater equal to a cubic spline function of time with each coefficient equal to a design parameter. This acts both to regularize the transient heater output and facilitates the use of design constraints that force the heater outputs to lie within a predefined range. The transient heater settings are optimized by minimizing a single objective function defined in terms of temperatures measured at discrete locations over the design surface and at various process times, which in turn are solved for using a variation of the infinitesimal-area analysis. Since the second-order objective function sensitivities are not readily available, minimization is carried out using a quasi-Newton method.

The second optimization methodology is for designing heater settings in low-thermal-inertia furnaces, where thermal radiation is the only mode of heat transfer and the design surface is the only enclosure surface that stores sensible energy. The problem is first linearized by specifying the temperature and the rate of temperature rise over the design surface, and the transient heater settings that provide the required sensible energy by uniformly heating the design surface are then found by carrying out an optimization at each time step. Each objective function is minimized using a quasi-Newton method, and an active set method is used to ensure a nonnegative heat flux distribution over the heater surface throughout the process. Since the underside of the design surface is insulated in most practical applications, it is necessary to verify the optimal heater settings by calculating the transient temperature distribution produced over the design surface when the adiabatic boundary condition is enforced.

It should also be noted that, although any one of these problems could be solved by coupling a generic analysis code directly to a commercially available optimization algorithm, this would be a computationally-expensive and time-consuming process because the objective sensitivities could only be found using a finite difference approximation. In contrast, the majority of methodologies presented in this dissertation take advantage of unique properties of the governing heat transfer equations to calculate the first-order (and sometimes second-order) sensitivities efficiently. In order to take advantage of the availability of the sensitivities, however, it is important to be judicious when selecting a minimization technique. In the methodologies presented in Chapters 2 and 3, for example, the second-order sensitivities used to form the Hessian are readily

available so Newton's method was chosen to carry out the minimization. In transient problems, on the other hand, second-order sensitivities are far more expensive to calculate, and because of this the objective function was minimized using the quasi-Newton method.

## 6.2 Recommendations for Future Work

### 6.2.1 Application of MINLP and Metaheuristic Minimization Techniques

In all of the design methodologies presented in this dissertation, the number of heaters used to heat the design surface was specified and their heat outputs were controlled by the design parameters. Since the heat generated by a particular heater can assume any value within its operating range, these design parameters are continuously variable and nonlinear programming techniques such as Newton's method can be used to minimize the objective function.

Suppose that the number and sizes of the heaters located on the heater surface are variables in the design problem and that the heaters are only available in several different sizes; such is the case in the design problem shown in Fig. 6.1. Because the design parameters corresponding to these variables hold discrete or integer values, nonlinear programming techniques cannot be used to optimize these parameters. One way to solve this problem is to find the optimum heater settings using gradient-based minimization for every possible combination of heaters, and then identifying the combination that produces the smallest value of  $F(\Phi^*)$ . While this may be reasonable in problems involving only a few heaters, this procedure becomes prohibitively time-consuming as the number of possible heater configurations increases.

An alternative strategy is to solve the design problems using a mixed-integer nonlinear programming (MINLP) technique. These techniques work by repeatedly solving for local minima over multiple subspaces of the feasible region in which all the design parameters are continuously variable using nonlinear programming techniques,

while a deterministic algorithm is used to identify subspaces that may contain the smallest local minima, so that not all of the subspaces need to be checked. For example, the BARON algorithm developed by Sahinidis (1996) uses a modified “branch-and-bound” technique to eliminate subspaces that could not contain the global minimum. Although these methods are guaranteed to identify the global minimum only if the feasible region of every subspace is known to be convex, a very good solution can usually be found even when this is not strictly true.

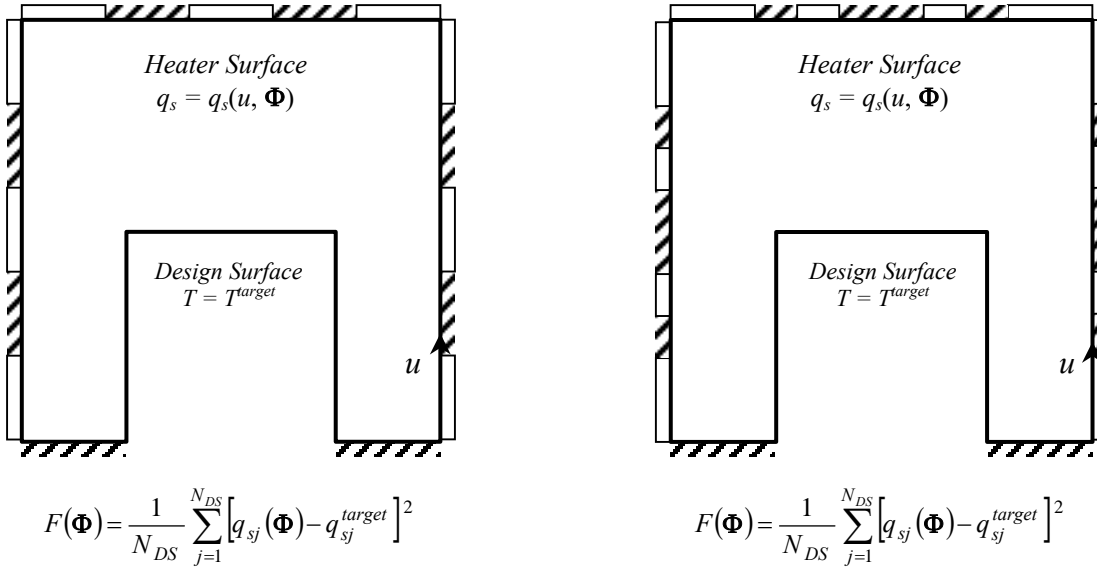


Figure 6.1: Two possible heater surface configurations for a radiant enclosure design problem where the number, size, and location of radiant heaters are design variables.

This problem could also be solved using a metaheuristic technique. Unlike deterministic optimization methods, metaheuristic techniques solve the optimization problem “approximately,” in the sense that the performance guarantees of the NLP and

MINLP algorithms are relaxed in order to increase computational efficiency. These techniques are particularly useful for solving combinatorial “traveling salesman” type problems and multimodal problems in which the objective function has many local minima over the feasible region; the computational effort required to solve these problems using a deterministic approach would be prohibitive.

Almost all metaheuristic algorithms work by comparing the objective functions corresponding to different sets of design parameters that are initially selected at random over the feasible region. As the method progresses, the selection process increasingly favors designs that produce small objective functions, while other designs are discarded; this procedure continues until the improvement made in successive trials becomes smaller than some user-defined criteria, at which point the algorithm is terminated. Two popular metaheuristic techniques are genetic algorithms (Forrest, 1993) and simulated annealing (Corana et al., 1987). Both of these methods are based on phenomena in the natural world; the former mimics evolutionary natural selection within species, while the latter is inspired by the Brownian motion of atoms that allow them to find positions within a crystal lattice that correspond with the lowest possible energy state when metals are slowly cooled in an annealing process. While they typically require more objective function evaluations than gradient-based methods to solve problems in which most of the design parameters are continuously variable and the objective function has few local minima, metaheuristic methods are often superior when many of the design parameters hold discrete or integer values and when many local minima exist over the feasible region. Furthermore, since the objective function sensitivities are not used to update the

design parameters, this method does not require the objective function to be smooth over the feasible region. Thus, unlike gradient-based techniques, metaheuristics can be used to optimize the geometry of diffuse-walled obstructed enclosures.

#### 6.2.2 Development of Optimization Methodologies to Design Enclosures Containing Participating Media

Although gradient-based optimization techniques have been used to solve many different types of design problems involving radiant enclosures, they have yet to be applied to design radiant enclosures containing participating media. These problems are encountered in diverse industrial settings, including glass and polymer fabrication, drying applications, and combustion processes. An example problem is shown in Fig. 6.2, where the objective is to identify the location and heat output of burners that produce the desired heat flux and temperature distribution over the design surface.

Recently, there has been a great interest in developing optimization methodologies to design combustion systems, with the objective of improving fuel efficiency and reducing pollution. Several preliminary optimization studies have identified designs that perform considerably better than existing designs originally found by trial-and-error. For example, Chen and Ryoji (2000) designed a post-flame reactor to reduce the amount of  $NO_x$  generated by a combustion process and Hiroyasu et al. (2003) optimized the injection system of a Diesel engine. The majority of these preliminary studies rely on genetic algorithms to carry out the minimization; since these algorithms require the evaluation of many prospective designs to solve the design problem, however,



limited computational resources restrict the system models to be rudimentary and largely phenomenological. Since a gradient-based optimization methodology would likely require fewer objective function evaluations to solve the design problem, it would then be possible to accommodate more sophisticated and accurate system model.

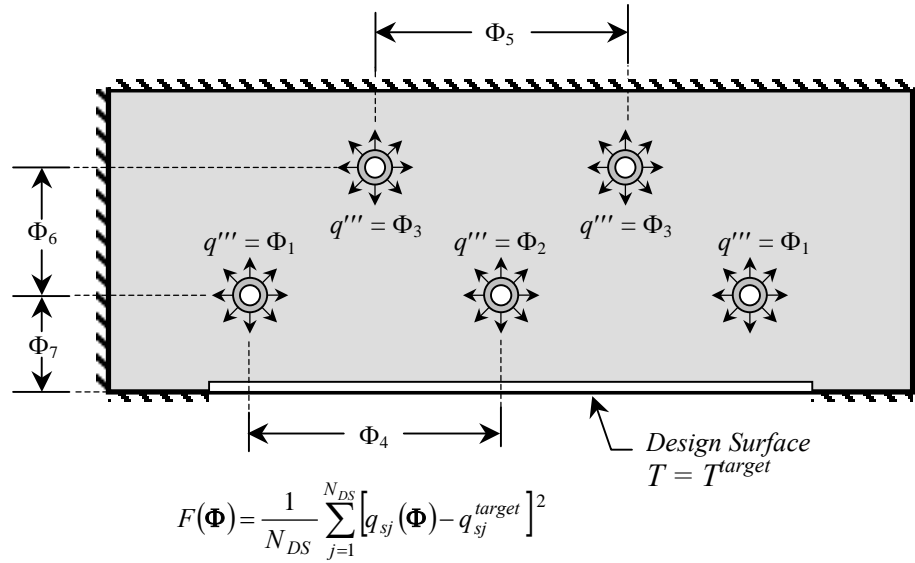


Figure 6.2: Example design problem involving a radiant enclosure containing a participating medium. Design parameters control the heat output and location of multiple burners located within a participating medium.

### 6.2.3 Integration of a CFD Code into the Optimization Design Methodology

In the majority of radiant enclosure design problems in the literature, convection heat transfer between the enclosure walls and a surrounding fluid are treated by specifying a convection coefficient over the surfaces; the most sophisticated treatment of mixed radiation/convection heat transfer in a radiant enclosure design problem to date was done

by França et al. (2001), who solved for convection heat transfer between the enclosure walls and the medium flowing between the walls by first specifying the velocity distribution within the fluid. Many radiant enclosure design problems require a more sophisticated treatment of convection heat transfer, however. In problems involving natural convection, for example, the equations governing the heat transfer and fluid flow within the radiant enclosure are coupled and must be solved simultaneously.

One way of treating this type of design problem is to integrate a CFD code like FLUENT or TASCFLOW into the optimization methodology. This would be relatively easy to do since the optimization methodology solves the design problem in its “natural” implicit form, which facilitates the use of complex heat transfer models. In the radiant enclosure design problem, a gradient-based algorithm could use a CFD code to calculate the temperature or heat flux distribution and the corresponding sensitivities over the design surface, which in turn would be used to form the objective function and gradient vector. Commercial codes have already been used to solve other types of optimization design problems; for example, Ebrahimi et al. (1997) optimized the riser geometry of an investment casting application with a quasi-Newton method that incorporated FIDAP to calculate the temperature distribution and temperature sensitivities within the molten metal, which were then used to form the objective function and gradient vector at each step. In another study, Han et al. (2003) optimized the scroll geometry of an axial impeller in an air conditioner, with the objective of achieving the maximum volumetric flow rate through the fan. The objective function was evaluated at various locations over the design surface by performing a CFD analysis, and these discrete objective function

values were then used to fit a quadratic response surface that approximated a continuously-variable objective function. The optimal scroll geometry corresponds with the global minimum of the response surface, which was found using Newton's method.

It should be noted, however, that the time and computational effort needed to carry out an analysis increases with the complexity of the heat transfer model. This is especially true in optimization design problems, since the existence of an efficient way to calculate the system response sensitivities becomes less likely as the model becomes more complex. In the most extreme case, the system response sensitivities are unavailable and the gradient vector must either be approximated by a finite difference technique or through the response surface method, which both require many objective function evaluations and a substantial amount of computational effort to carry out.

## **Appendix A:**

# **Optimization through Nonlinear Programming**

## A.1 Introduction

Optimization techniques are encountered in many diverse settings; most often, as the name implies, their purpose is to determine the configuration of a system that results in optimal performance. For example, they may be used to determine a financial portfolio that maximizes profit made on the stock market over a set time period, the number of trucks a hauling company should dispatch to make deliveries to different cities in order to minimize operating cost, or the geometry of a rocket nosecone that minimizes drag and heating due to thermal dissipation.

The first step in solving an optimization problem is to restate it as a minimization problem. In particular, it is necessary to designate a set of *design parameters* contained in the vector  $\Phi$  that control the design configuration, and then to define an *objective function*,  $F(\Phi)$ , which quantifies the “goodness” of the configuration corresponding to the set of design parameters contained in  $\Phi$ . The objective function is usually defined in such a way that the minimum of  $F(\Phi)$  corresponds to the ideal design outcome. For example, in the hauling problem described in the previous paragraph, a suitable objective function would be the overall cost associated with making the deliveries including driver time, fuel charges, and truck maintenance and depreciation, while the design parameters would define the number of trucks dispatched to make the deliveries and the sequence in which the deliveries are made. Thus, the optimal design configuration corresponds to the design parameters contained  $\Phi^*$ , which in turn is found by minimizing  $F(\Phi)$ ,

$$F(\Phi^*) = \text{Min}[F(\Phi)], \Phi \in \mathcal{R}^n. \quad (\text{A.1})$$

Usually, it is also necessary to impose *design constraints* having the form

$$c_i(\Phi) = 0, \quad i = 1 \dots m', \quad (\text{A.2})$$

and

$$c_i(\Phi) \geq 0, \quad i = m'+1 \dots m, \quad (\text{A.3})$$

which define the domain (or feasible region) of  $\Phi$ . In the hauling problem, the design constraints would ensure that the optimum number of trucks is greater than zero but no more than the total number of trucks available in the company fleet.

If all the functions in Eqs. (A.1) through (A.3) are linear, then the optimal solution can be solved using linear programming techniques; if the constraints can be written as equalities, this is tantamount to solving a system of linear equations. In most situations, however, the objective function and/or one or more of the constraints will be a nonlinear function of the design parameters in which case nonlinear programming techniques must be used to identify  $\Phi^*$ .

At this point, it is convenient to designate two classes of nonlinear programming problems; those that only involve design parameters that continuously vary over the feasible region, and those that contain one or more design parameters that can only assume discrete values. For example, when determining the optimal heater settings of a process furnace, the heat flux produced by a heater can assume any value that lies within a designated minimum and maximum heater output, so the design parameters can vary continuously over their corresponding feasible region. In the hauling problem, on the other hand, the design parameters that designate the number of dispatched trucks and the order in which the deliveries are made can only assume integer values. Problems belonging to the first class are generally easier to solve, while those belonging to the

second class, called “mixed-integer nonlinear programming” problems, are more complex and require specialized techniques in order to find their solutions. Further discussion of nonlinear programming techniques is restricted to those that pertain to problems of the first class.

In this appendix, unconstrained gradient-based minimization techniques are first discussed and demonstrated in Section A.2. These techniques are modified to allow the treatment of some simple linear design constraints in Section A.3, and conclusions are presented in Section A.4.

## A.2 Unconstrained Optimization Techniques

Many different techniques have been developed to minimize the objective function. Gradient-based techniques are most commonly used when the objective function and design constraints are continuously differentiable over the feasible region. These techniques minimize  $F(\Phi)$  iteratively and work according to the following steps: at the  $k^{th}$  iteration,

1. The magnitude of the gradient vector is first checked to see if  $|\nabla F(\Phi^k)| \approx 0$ , in which case  $\Phi^k$  is a stationary point and the process is stopped. If the Hessian matrix is positive-definite at the stationary point (i.e. if  $\mathbf{d}^T \nabla^2 F(\Phi^k) \mathbf{d} > 0$  for any vector  $\mathbf{d}$ ) then  $\Phi^k$  is a local minimizer of  $F(\Phi)$ .
2. The search direction,  $\mathbf{p}^k$ , is chosen based on the local objective function curvature at  $\Phi^k$ .

3. A step size,  $\alpha^k$ , is selected. This is usually done by minimizing the function

$$f(\alpha^k) = F(\Phi^k + \mathbf{p}^k \alpha^k) \text{ with respect to } \alpha^k.$$

4. A step is taken in the  $\mathbf{p}^k$  direction,  $\Phi^{k+1} = \Phi^k + \mathbf{p}^k \alpha^k$ .

The computational efficiency of the algorithm is a function of the total number of iterations required to arrive at  $\Phi^*$  and the CPU time and memory required to perform each iteration. These factors, in turn, depend on how the search direction and step size are chosen and the computational effort required for solving the first- and second- order objective function sensitivities.

#### A.2.1 Choosing a Search Direction

The main difference between gradient-based optimization techniques lies in how the search direction is chosen. The search direction is chosen based on the first-order curvature contained in the gradient vector,

$$\nabla F(\Phi) = \left\{ \frac{\partial F(\Phi)}{\partial \Phi_1}, \frac{\partial F(\Phi)}{\partial \Phi_2}, \dots, \frac{\partial F(\Phi)}{\partial \Phi_n} \right\}^T, \quad (\text{A.4})$$

and in some cases, the second-order curvature information contained in the Hessian matrix,

$$\nabla^2 F(\Phi) = \begin{bmatrix} \partial^2 F / \partial \Phi_1 \partial \Phi_1 & \dots & \partial^2 F / \partial \Phi_1 \partial \Phi_n \\ \partial^2 F / \partial \Phi_2 \partial \Phi_1 & \partial^2 F / \partial \Phi_2 \partial \Phi_2 & \\ \vdots & & \ddots \\ \partial^2 F / \partial \Phi_n \partial \Phi_1 & & \partial^2 F / \partial \Phi_n \partial \Phi_n \end{bmatrix} \quad (\text{A.5})$$

When selecting a search direction, it is important to ensure that the choice of  $\mathbf{p}^k$  reduces the objective function, i.e.  $F(\Phi^k + \mathbf{p}^k \alpha^k) < F(\Phi^k)$  with  $\alpha^k > 0$ . This condition is satisfied as long as  $\mathbf{p}^k$  is a *descent direction*,



$$\mathbf{p}^{kT} \nabla F(\Phi_k) < 0. \quad (\text{A.6})$$

In gradient-based methods, the search direction is found by solving a system of linear equations,

$$\mathbf{A} \mathbf{p}^k = -\nabla F(\Phi_k), \quad (\text{A.7})$$

where the coefficient matrix,  $\mathbf{A}$ , differentiates the type of method used. It can be shown that  $\mathbf{p}^k$  is guaranteed to be a descent direction as long as the coefficient matrix is positive definite, i.e.  $\mathbf{d}^T \mathbf{A} \mathbf{d} > 0$  for any vector  $\mathbf{d}$ .

The most intuitive choice for a search direction is found by setting  $\mathbf{A}$  in Eq. (A.7) equal the identity matrix,

$$\mathbf{p}^k = -\nabla F(\Phi_k), \quad (\text{A.8})$$

which is the direction of steepest descent; accordingly, this method is called the “steepest descent” method. In this method,  $\mathbf{p}^k$  is always guaranteed to be a descent direction, and it only requires evaluation of first-order sensitivities. Nevertheless, the steepest descent method usually has a linear or sublinear rate of convergence towards  $\Phi^*$ , and therefore is rarely used as an optimization tool.

A much better search direction is found by considering both the first- and second-order objective function sensitivities. In Newton’s method, the search direction is derived from a second-order Taylor series expansion of the objective function about  $\Phi^k$ ,

$$F(\Phi^k + \mathbf{p}^k) = F(\Phi^k) + \mathbf{p}^{kT} \nabla F(\Phi^k) + \frac{1}{2} \mathbf{p}^{kT} \nabla^2 F(\xi^k) \mathbf{p}^k, \quad (\text{A.9})$$

where  $\xi$  is an undetermined value lying between  $\Phi^k$  and  $\Phi^k + \mathbf{p}^k$ . Let  $\Phi^* = \Phi^k + \mathbf{p}^k$ , and furthermore assume that the objective function is quadratic in which case the Hessian matrix is constant. (This assumption is usually reasonable in the vicinity of  $\Phi^*$ , where

the higher-order terms in the Taylor series approximation become very small.) The gradient vector at  $\Phi^* = \Phi^k + p^k$  can be found by taking the derivative of Eq. (A.9),

$$\nabla F(\Phi^*) \approx \nabla F(\Phi^k) + \nabla^2 F(\Phi^k) p^k. \quad (\text{A.10})$$

Since  $|\nabla F(\Phi^*)| = 0$ , Eq. (A.10) can be rearranged to solve for the search direction,

$$\nabla^2 F(\Phi^k) p^k = -\nabla F(\Phi^k). \quad (\text{A.11})$$

Here,  $p^k$  is Newton's direction, and hence this method is referred to as "Newton's method." This method usually requires the fewest iterations to reach  $\Phi^*$ , and is used as a benchmark to measure the performance of other techniques. The drawback to this approach, however, is that both the first- and second-order objective function sensitivities are required to solve for  $p^k$ . Thus, the viability of Newton's method depends on the computational effort required to calculate the Hessian matrix. If the second-order objective function sensitivities are found efficiently (which is the case for most diffuse-walled radiant enclosure problems) then Newton's method is usually the most efficient optimization technique. More often, however, significant computational effort is required to calculate the Hessian matrix. In the worst case, the first- and second-order sensitivities are found using a forward difference approximation, which requires  $n$  objective function evaluations to estimate  $\nabla F(\Phi^k)$  and an extra  $(n+1) \times n/2$  objective function evaluations to estimate  $\nabla^2 F(\Phi^k)$ . In these cases, the advantage of requiring fewer iterations to find  $\Phi^*$  is usually negated by the computational effort required to solve the Hessian at each iteration.

The quasi-Newton method avoids the computational effort associated with calculating the second-order objective function sensitivities by using first-order curvature

information collected from previous iterations to estimate the Hessian matrix. In this method, the search direction is found by solving

$$\mathbf{B}^k \mathbf{p}^k = -\nabla F(\Phi^k), \quad (\text{A.12})$$

where  $\mathbf{B}^k$  is an approximation matrix. Initially,  $\mathbf{B}^0$  is set equal to the identity matrix, and  $\mathbf{p}^k$  is then the steepest-descent direction. The Hessian approximation is improved in subsequent iterations by adding an update matrix,

$$\mathbf{B}^{k+1} = \mathbf{B}^k + \mathbf{U}^k, \quad (\text{A.13})$$

where  $\mathbf{U}^k$  is calculated using values of the objective function and gradient vector calculated from previous iterations. The Hessian is accurately approximated after few iterations, and  $\mathbf{B}^k$  converges to  $\nabla^2 F(\Phi^k)$  as  $k$  becomes large.

Many different update schemes have been proposed; the most popular is the Broyden-Fletcher-Goldfarb-Shanno (BFGS) scheme (Gill et al., 1986a). Here, the update matrix is given by

$$\mathbf{U}_k = -\frac{\mathbf{B}^k \mathbf{s}^k \mathbf{s}^{kT} \mathbf{B}^k}{\mathbf{s}^{kT} \mathbf{B}^k \mathbf{s}^k} + \frac{\mathbf{y}^k \mathbf{y}^{kT}}{\mathbf{y}^{kT} \mathbf{s}^k}, \quad (\text{A.14})$$

where  $\mathbf{s}^k = \Phi^{k+1} - \Phi^k$  and  $\mathbf{y}^k = \nabla F(\Phi^{k+1}) - \nabla F(\Phi^k)$ . Equation (A.14) is simplified by substituting in Eq. (A.12) and noting that  $\mathbf{p}^k \alpha^k = \mathbf{s}^k$ ,

$$\mathbf{U}_k = -\frac{\nabla F^T(\Phi^k) \nabla F(\Phi^k)}{\nabla F^T(\Phi^k) \mathbf{p}^k} + \frac{\mathbf{y}^k \mathbf{y}^{kT}}{\alpha^k \mathbf{y}^{kT} \mathbf{p}^k}. \quad (\text{A.15})$$

Because  $\mathbf{p}^k$  is initially equal to the steepest-descent direction and accurately approximates Newton's direction only after several iterations, the quasi-Newton's method usually requires more iterations to identify  $\Phi^*$  than Newton's method. Since the second-order sensitivities are not calculated, however, the computational effort required to minimize  $F(\Phi)$  is often much less than that required by Newton's method.

It should be noted that while  $\mathbf{p}^k$  is always guaranteed to be a descent direction in the steepest-descent method, this is not the case for the Newton and quasi-Newton methods. In particular, both Eqs. (A.11) and (A.12) will fail to produce a descent direction if  $\nabla^2 F(\Phi^k)$  is semi-definite, which happens if there exists some vector  $\mathbf{d}$  that satisfies  $\mathbf{d}^T \nabla^2 F(\Phi^k) \mathbf{d}^T = 0$  with  $|\mathbf{d}| \neq 0$ ; such a condition arises when two or more of the rows in the Hessian matrix are linearly dependent. The Newton and quasi-Newton methods will also fail if  $\nabla^2 F(\Phi^k)$  is negative-definite, where there exists some vector  $\mathbf{d}$  that satisfies  $\mathbf{d}^T \nabla^2 F(\Phi^k) \mathbf{d}^T < 0$ , even if there exists a descent direction at  $\Phi^k$ .

This situation is most easily visualized by considering the case where Newton's method in one-dimension,

$$x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}, \quad (\text{A.16})$$

is applied to minimize the function shown in Fig. A.1. (Note that this equation is analogous to Eq. (A.11).) In the immediate vicinity of  $x^*$ ,  $a < x^k < b$ ,  $f(x^k)$  is concave so  $f''(x^k)$  is greater than zero; applying Eq. (A.16) in this region would produce a point  $x^{k+1}$  where  $f(x^{k+1}) < f(x^k)$  and  $x^*$  would be reached after only a few iterations. For  $b < x^k < c$ , however,  $f(x^k)$  is linear and  $f''(x^k) = 0$ ; applying Eq. (A.16) in this region would result in a division-by-zero error. Finally,  $f(x^k)$  is concave for  $x^k > c$ , so  $f'(x^k) > 0$  and  $f''(x^k) < 0$ ; applying Eq. (A.16) would produce a point  $x^{k+1}$  to the right of  $x^k$  where  $f(x^{k+1}) > f(x^k)$ , and the algorithm would eventually reach the local maximum at  $x^{*2}$ .

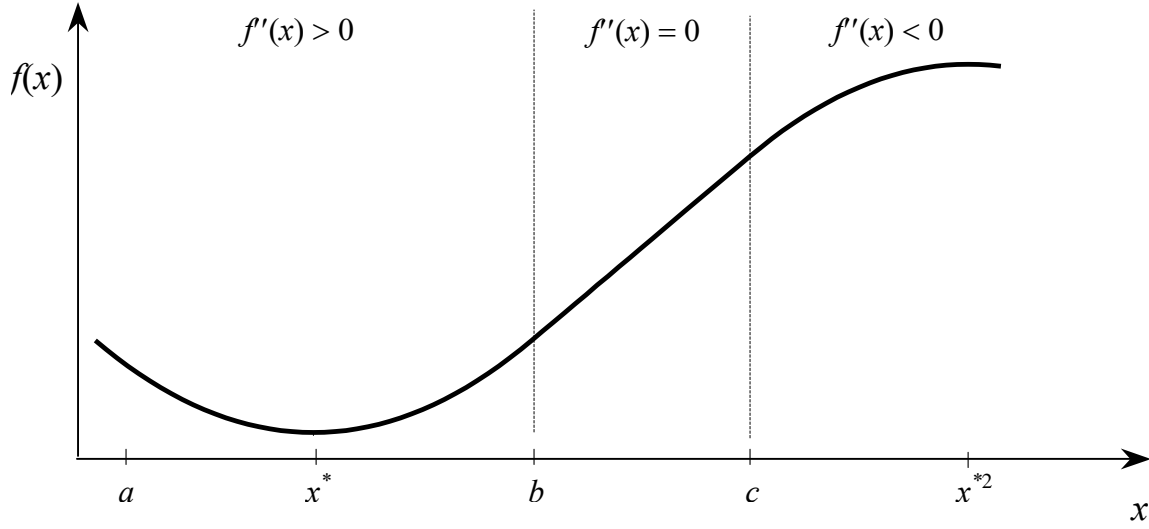


Figure A.1: Newton's method in one dimension.

To understand this phenomenon in multiple dimensions, it is useful to first decompose the Hessian into its eigenvalues,  $\{\lambda_i, i = 1 \dots n\}$ , and eigenvectors,  $\{\mu_i, i = 1 \dots n\}$ . In particular, the  $j^{th}$  eigenvalue and eigenvector of  $\nabla^2 F(\Phi^k)$  satisfy

$$\nabla^2 F(\Phi^k) \mu_j = \lambda_j \mu_j. \quad (\text{A.16})$$

If  $\lambda_j$  is equal to zero at  $\Phi^k$ , then  $F(\Phi^k)$  becomes a linear function in the  $\mu_j$  direction. In this case both  $\nabla^2 F(\Phi^k)$  and  $\mathbf{B}^k$  are singular so Eqs. (A.11) and (A.12) cannot be solved. Likewise, if  $\lambda_j < 0$ , then  $F(\Phi^k)$  is concave in the  $\mu_j$  direction and Eqs. (A.11) and (A.12) are no longer guaranteed to produce a descent direction, even though one may exist.

For example, consider the quadratic objective function given by

$$F(\Phi) = F(x, y) = 4x^2 + y^2 + 4xy - 3x - 2y, \quad (\text{A.17})$$

which is to be minimized starting from  $\Phi^0 = \{-2, -2\}^T$ . By differentiating  $F(\Phi)$  with respect to  $x$  and  $y$ , the gradient vector is found to be

$$\nabla F(x, y) = \{8x + 4y - 3, 2y + 4x - 2\}^T, \quad (\text{A.18})$$

and the Hessian matrix is given by

$$\nabla^2 F = \begin{bmatrix} 8 & 4 \\ 4 & 2 \end{bmatrix}, \quad (\text{A.19})$$

with eigenvalues and eigenvectors equal to  $\lambda_1 = 0$ ,  $\mu_1 = \{-1, 2\}^T$ ,  $\lambda_2 = 10$ , and  $\mu_2 = \{2, 1\}^T$ . Because the rows of the Hessian matrix are linearly dependent, Eq. (A.11) will not produce a descent direction even though one must exist at  $\Phi^0$ , since  $\nabla F(\Phi^0) = \{-27, -14\}^T$ . As shown graphically in Fig. A.2, the null eigenvalue indicates that  $F(\Phi)$  is linear in the  $\mu_1$  direction.

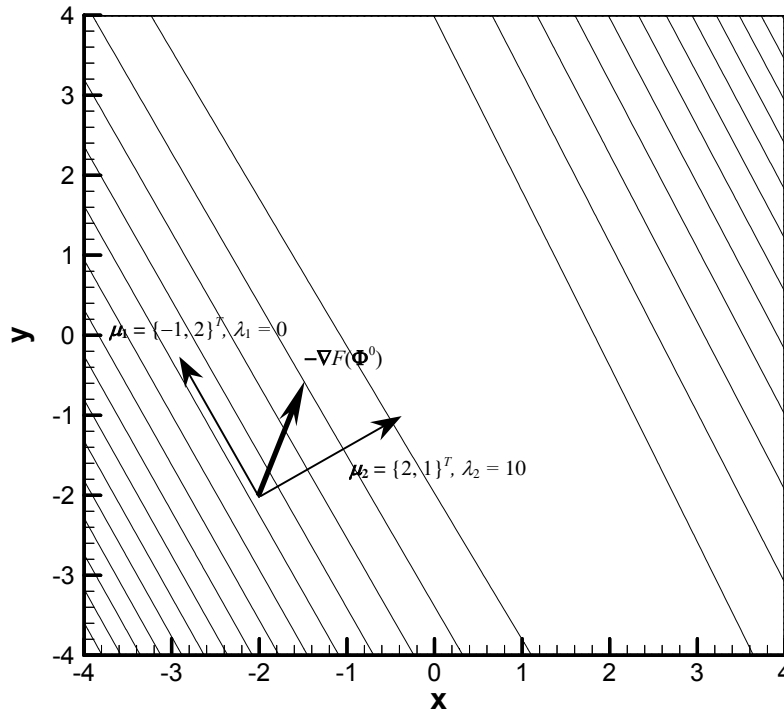


Figure A.2: An example of an objective function with an indefinite Hessian.

These situations can be diagnosed and corrected by first performing a Cholesky decomposition on the Hessian,

$$\nabla^2 F(\Phi^k) = \mathbf{L} \mathbf{D} \mathbf{L}^T, \quad (\text{A.20})$$

or approximate Hessian matrices, and by recalling that the back-substitution step used to solve a system of linear equations involves division by the diagonal elements of  $\mathbf{D}$ . If  $\nabla^2 F(\Phi^k)$  or  $\mathbf{B}^k$  is positive definite, the diagonal elements of  $\mathbf{D}$  will be strictly positive, and the search direction can be solved by substituting Eq. (A.20) into Eq. (A.11) and solving by back-substitution. If the Hessian is indefinite, however, at least one of the diagonal elements in  $\mathbf{D}$  equals zero and the back-substitution step results in a division-by-zero error. Finally, if one of the diagonal elements of  $\mathbf{D}$  is negative, then the original matrix is negative definite, and a descent direction is no longer guaranteed.

This situation is most often resolved by adding a diagonal matrix  $\mathbf{E}^k$  to an indefinite or negative-definite Hessian to produce a related positive-definite matrix,  $\tilde{\nabla}^2 F(\Phi^k)$ ,

$$\tilde{\nabla}^2 F(\Phi^k) = \nabla^2 F(\Phi^k) + \mathbf{E}^k, \quad (\text{A.21})$$

where the elements in  $\mathbf{E}$  are sufficiently large to ensure that the diagonal elements of  $\mathbf{D}$  are strictly non-negative, yet small enough so as to retain an accurate approximation of  $\nabla^2 F(\Phi)$ . (This is, in fact, a *regularization* technique, similar to the ones described in Appendix B.) There are many different strategies for choosing  $\mathbf{E}$ , including those discussed in Gill et al. (1986b) and Betsekas (1999b). Both of these approaches are also used to resolve indeterminate Hessian approximations encountered with the quasi-Newton method.

In the methods discussed so far,  $\mathbf{p}^k$  is chosen based solely on the objective function curvature at  $\Phi^k$  without considering search directions used at previous iterations. Because of this, there is no guarantee that progress made by minimizing in one search

direction will not be “spoiled” by minimizations in subsequent directions. In an extreme example, consider the minimization of the quadratic objective function shown in Fig. A.3, where successive search directions are alternately set equal to the unit vectors  $\hat{e}_x$  and  $\hat{e}_y$  in the  $x$ - and  $y$ - coordinate directions, respectively. When minimization has been carried out in the  $\hat{e}_x$  direction at the  $k^{th}$  iteration, say, then  $y^{k-1}$  no longer minimizes  $F(x^k, y)$ , and an updated value of  $y$  must be found in the next iteration.

Another class of gradient-based minimization techniques, called conjugate gradient techniques, avoid this problem by finding a set of “non-interfering” (or mutually conjugate) search directions. These were originally developed to solve systems of  $n$  equations containing  $n$  unknowns,

$$\mathbf{A}\Phi = \mathbf{b}, \quad (\text{A.22})$$

which is equivalent to minimizing the quadratic objective function

$$F(\Phi) = \frac{1}{2} \Phi^T \mathbf{A} \Phi - \mathbf{b}^T \Phi, \quad (\text{A.23})$$

where  $\nabla F(\Phi) = \mathbf{A} \Phi - \mathbf{b}$  and  $\nabla^2 F(\Phi) = \mathbf{A}$ . The solution strategy is to first identify a set of  $n$  directions that are mutually conjugate with respect to the  $\mathbf{A}$  matrix,  $\{\mathbf{p}^i, i = 1 \dots n\}$ , which by definition satisfy

$$\mathbf{p}^{iT} \mathbf{A} \mathbf{p}^j = 0, \quad i \neq j. \quad (\text{A.24})$$

(If  $\mathbf{A}$  is the identity matrix, then  $\mathbf{p}^i$  and  $\mathbf{p}^j$  are orthogonal for  $i \neq j$ .) The conjugate set is generated using the residual vector  $\mathbf{r}^i = \mathbf{b} - \mathbf{A}\Phi^i$ , through the recursion (Nash and Sofer, 1996)

$$\mathbf{p}^k = -\mathbf{r}^k + \beta^{k-1} \mathbf{p}^{k-1}, \quad (\text{A.25})$$



$$\alpha^k = \frac{\|\mathbf{r}^k\|_2^2}{\mathbf{p}^{kT} \mathbf{A} \mathbf{p}^k}, \quad (\text{A.26})$$

$$\Phi^{k+1} = \Phi^k + \alpha^k \mathbf{p}^k, \quad (\text{A.27})$$

$$\mathbf{r}^{k+1} = \mathbf{r}^k + \alpha^k \mathbf{A} \mathbf{p}^k, \quad (\text{A.28})$$

and

$$\beta^k = \frac{\|\mathbf{r}^{k+1}\|_2^2}{\|\mathbf{r}^k\|_2^2}, \quad (\text{A.29})$$

with the initial values  $\mathbf{r}^0 = \mathbf{b} - \mathbf{A}\Phi^0$ ,  $\mathbf{p}^{-1} = 0$ , and  $\beta^0 = 0$ . If exact algebra is used, then Eqs. (A.25) – (A.29) find  $\Phi^*$  in exactly  $n$  iterations. (The matrix  $\mathbf{A}$  must be symmetric and positive definite, however, for the reasons discussed previously in relation to the Newton and quasi-Newton methods.)

The above method can be extended to minimize objective functions of arbitrary form. By applying identities based on conjugacy to the above recursion, the search directions are generated using only the gradient vector,

$$\beta^k = \frac{\nabla^T f(\Phi^k) \nabla f(\Phi^k)}{\nabla^T f(\Phi^{k-1}) \nabla f(\Phi^{k-1})}, \quad (\text{A.30})$$

and

$$\mathbf{p}^k = -\nabla f(\Phi^k) + \beta^k \mathbf{p}^{k-1}, \quad (\text{A.31})$$

thus avoiding the computational effort required to calculate the Hessian matrix at each iteration. (The step size is found by performing a line search from  $\Phi^k$  in the  $\mathbf{p}^k$  direction.) Since the objective function is not generally quadratic, the set of search directions will not satisfy the conjugacy conditions in Eq. (A.24) exactly and thus more than  $n$  steps are usually needed to find  $\Phi^*$ . Nevertheless,  $F(\Phi)$  is usually approximated as a quadratic in

the vicinity of  $\Phi^*$  with sufficient accuracy in order to make this an efficient minimization technique.

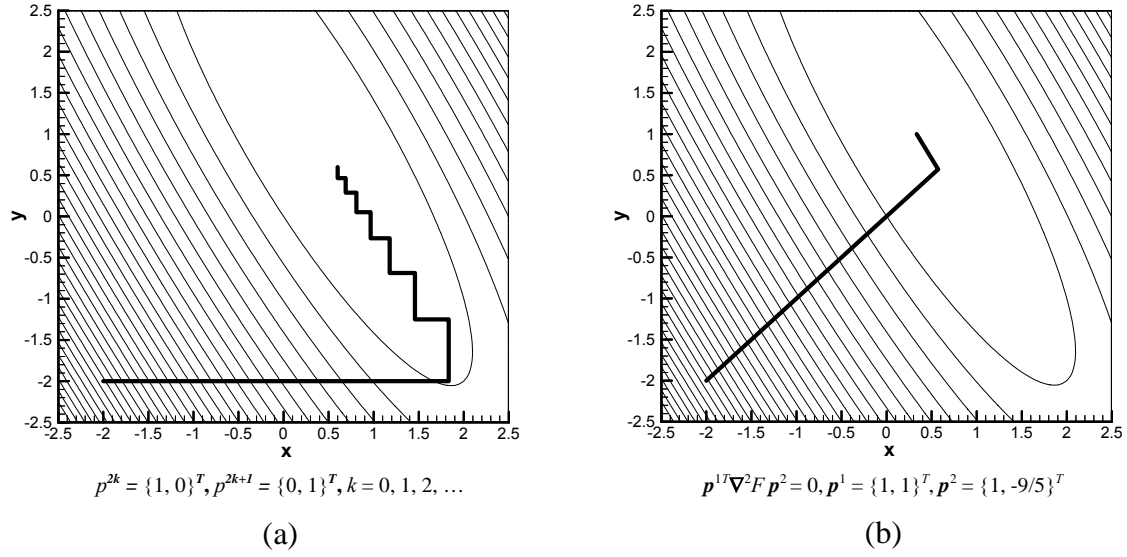


Figure A.3: Minimization of  $F(x, y) = 4x^2 + y^2 + 3xy - 5x - 3y$  using (a) the coordinate directions, and (b) conjugate vectors as search directions.

### A.2.2 Choosing a Step Size

The next task is to choose a step size,  $\alpha^k$ , which determines how far the design parameters should be moved from  $\Phi^k$  in the direction  $\mathbf{p}^k$  at the  $k^{th}$  iteration. Many different schemes have been developed for choosing the step size. As is the case when selecting a search direction method, the objective when selecting a step size scheme is to choose the one that minimizes the overall computational effort required to find  $\Phi^*$ ; accordingly, this decision is based on both the computational effort required to calculate the objective function and the availability of the first- and second-order objective function sensitivities.

The most intuitive strategy is to use the step size that maximizes the decrease in the objective function at that particular iteration,  $F(\Phi^k) - F(\Phi^{k+1})$ . As shown in Fig. A.4, this is equivalent to solving the univariate minimization problem

$$f(\alpha^{k*}) = \text{Min}_{\alpha^k > 0} [f(\alpha^k)], \quad (\text{A.32})$$

where  $f(\alpha^k) = F(\Phi^k + \mathbf{p}^k \alpha^k)$ . This strategy is called an “exact line search,” and usually requires the fewest number of iterations to find  $\Phi^*$  when implemented with a given search direction scheme.

Nevertheless, the extra computational effort needed at each iteration to find the exact value of  $\alpha^{k*}$  usually exceeds the amount saved by reducing the total number of iterations. Because of this, in most practical implementations the minimization procedure is stopped once a step size is found that approximates  $\alpha^{k*}$  within a specified tolerance. One common stopping criterion is the Armijo rule (Bertsekas, 1999a), which requires the magnitude of the objective function decrease obtained with a given step size to be greater

than some specified fraction of the decrease predicted by a first-order Taylor series approximation of the objective function,

$$\left| F(\Phi^k) - F(\Phi^k + \mathbf{p}^k \alpha^k) \right| \geq \left| \mu \alpha^k \mathbf{p}^{kT} \nabla F(\Phi^k) \right|, \quad (\text{A.33})$$

where  $\mu$  is the Armijo parameter, which typically has a value between  $10^{-5} \leq \mu \leq 10^{-2}$  (Bertsekas, 1999a). The Armijo rule ensures that the step size is small enough so that the 2<sup>nd</sup>-order Taylor series expansion of the objective function about  $\Phi^k$  used to calculate  $\mathbf{p}^k$  accurately models the objective function at  $\Phi^{k+1}$ .

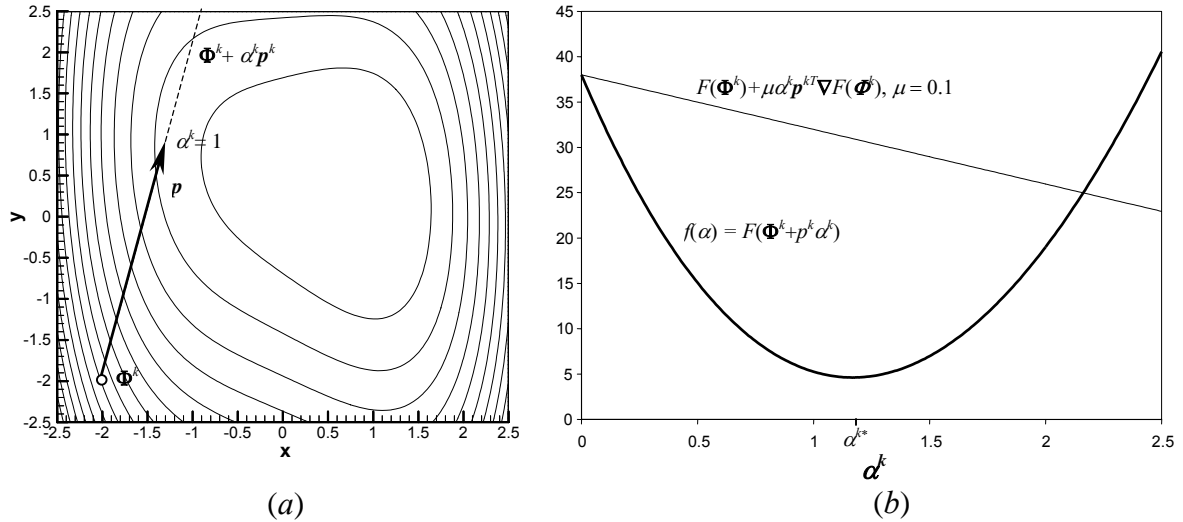


Figure A.4: Step size selection as a univariate minimization problem; (a) contour plot of  $F(\Phi) = F(x, y) = x^4 + 2y^2 + xy - 3x - 2y$  with  $\Phi^k = \{-2, -2\}^T$  and  $\mathbf{p}^k = \{0.71, 2.82\}^T$  (Newton's direction), and (b) univariate minimization of  $f(\alpha^k) = F(\Phi^k + \alpha^k \mathbf{p}^k)$  subject to the Armijo rule, with  $\mu = 10^{-1}$ .

There are many possible ways to perform a line search. If the first and second-order objective function sensitivities are readily available, then  $\alpha^{k*}$  (or its approximation)

can be found through a one-dimensional implementation of Newton's method. At the  $i^{th}$  sub-iteration, the new estimate for  $\alpha^{k*}$  is given by

$$\alpha^{k^{i+1}} = \alpha^{k^i} - \frac{f'(\alpha^{k^i})}{f''(\alpha^{k^i})}, \quad (\text{A.34})$$

where  $f'(\alpha^{k^i}) = \mathbf{p}^{k^T} \nabla F(\Phi^k + \mathbf{p}^k \alpha^{k^i})$  and  $f''(\alpha^{k^i}) = \mathbf{p}^{k^T} \nabla^2 F(\Phi^k + \mathbf{p}^k \alpha^{k^i}) \mathbf{p}^k$ . Just as in the multivariate case, this method typically requires the fewest number of subiterations since both first- and second-order curvature information is used to update the guess for  $\alpha^{k*}$ , although the computational effort needed to calculate the Hessian at each sub-iteration is usually prohibitive.

Quadratic and cubic interpolation are also used to select a near optimum step size. The procedure works as follows; at each subiteration, an interval is first selected over which  $\alpha^{k*}$  is known to lie, i.e.  $a^i \leq \alpha^{k*} \leq b^i$ . Next, either a quadratic or a cubic polynomial is fitted to values of  $f(\alpha^k)$  and  $f'(\alpha^k)$  within  $[a_i, b_i]$ , and an estimate of  $\alpha^{k*}$ ,  $\alpha^{ki*}$ , is then estimated from the polynomial minimum. Finally, a smaller updated interval is identified around  $\alpha^{ki*}$  and the interpolation process is repeated. Cubic interpolation usually provides a more accurate estimate of  $f(\alpha^k)$  and thus requires fewer iterations than quadratic interpolation to find an accurate approximation of  $\alpha^{k*}$ . On the other hand, whereas quadratic interpolation only uses evaluations of  $f(\alpha^k)$ , cubic interpolation is based on values of both  $f(\alpha^k)$  and  $f'(\alpha^k)$ , and therefore tends to require more computational effort than quadratic interpolation at each subiteration.

In many instances, the magnitude of the search direction,  $|\mathbf{p}^k|$ , accurately estimates the distance between  $\Phi^k$  and  $\Phi^k + \alpha^{k*} \mathbf{p}^k$ , so that  $\alpha^{k*} \approx 1$ . This is often the case when the

Newton or quasi-Newton search direction is used, since  $\mathbf{p}^k$  set equal to the distance between  $\Phi^k$  and  $\Phi^*$  in the second-order Taylor series approximation of  $F(\Phi)$  used to derive Eqs. (A.11) and (A.12). In this case, it is often sufficient to use a backtracking search to find a step size, where the search direction is set equal to  $2^{-m}$  with  $m$  being the first non-negative integer that satisfies the Armijo criterion.

Finally, if the objective function evaluations are computationally expensive, it is often impractical to use a line search to solve for  $\alpha^k$ . An alternative is to employ a diminishing step size rule, where  $\alpha^k$  is a specified function of  $k$  that satisfies the conditions

$$\lim_{k \rightarrow \infty} \alpha^k = 0, \quad (\text{A.35})$$

and

$$\sum_{k=1}^{\infty} \alpha^k = \infty. \quad (\text{A.36})$$

The second condition prevents the minimization process from converging to a non-stationary point. One common approach is to set the step size equal to a non-diminishing sequence based on iteration number,

$$\alpha^k = \frac{\alpha^0}{k^a}, \quad 0 < a \leq 1. \quad (\text{A.37})$$

Although the diminishing step size rule has some favorable convergence properties (Bertsekas, 1999b), convergence tends to be slow since  $\alpha^k$  is not always a good approximation of  $\alpha^{k*}$ . Accordingly, this method is usually restricted to situations when slow convergence is inevitable; for example, when an accurate estimator of  $\nabla F(\Phi^k)$  is unavailable.

### A.2.3 Demonstration of Unconstrained Minimization Techniques

The Newton, quasi-Newton (BFGS), steepest descent, and conjugate gradient minimization techniques discussed in the previous sections are demonstrated by applying them to minimize Rosenbrock's function,

$$F(x, y) = (x - 1)^2 + 10(x^2 - y)^2. \quad (\text{A.38})$$

This function is often used to evaluate gradient-based optimization techniques, because as shown in Fig. A.5, the narrow, twisting valley leading up to the global minimum at  $\Phi^* = \{1, 1\}^T$  makes it a difficult objective function to minimize.

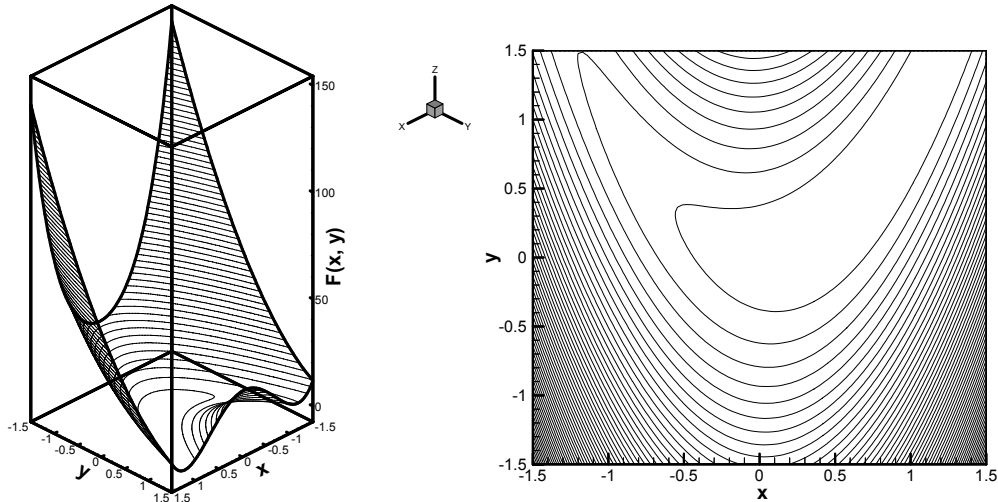


Figure A.5: Rosenbrock's Function,  $F(x, y) = (x - 1)^2 + 10(x^2 - y)^2$ .

In each case, minimization starts at  $\Phi_0 = \{1, -1\}^T$ , and is stopped when  $|\nabla F(\Phi^k)| < 10^{-6}$ . The step size at each iteration is found by performing repeated quadratic polynomial interpolations until the Armijo criterion is satisfied with  $\mu = 10^{-2}$ . The minimization paths are shown in Fig. A.6, while the decrease in the objective function with each successive iteration is plotted in Fig. A.7. Performance of the minimization

techniques is also summarized in Table A.1, and is measured based on both the total iterations and the number times  $F(\Phi)$ ,  $\nabla F(\Phi)$ , and  $\nabla^2 F(\Phi)$  need to be evaluated in order to find  $\Phi^*$ .

Each of the minimization techniques was able to find  $\Phi^*$ . As predicted, Newton's method requires the fewest iterations, since the search direction is chosen based on first- and second-order curvature information evaluated at each iteration. The steepest descent method, on the other hand, requires the most iterations because the search direction is chosen based only on the first-order curvature of  $F(\Phi)$ . Although the quasi-Newton method initially starts with the steepest descent direction, it required only seven more iterations than Newton's method to find  $\Phi^*$ , suggesting that an accurate approximation of the Hessian matrix was formed after only a few iterations.

If approximately the same computational effort is required to calculate  $F(\Phi)$ ,  $\partial F(\Phi)/\partial \Phi_p$ , and  $\partial^2 F(\Phi)/\partial \Phi_p \partial \Phi_q$  (as is usually the case when the sensitivities are found analytically) then Newton's method the most efficient of the four methods. This is usually only true, however, if the first- and second-order sensitivities can be calculated efficiently; otherwise, any advantage gained by performing fewer iterations to reach  $\Phi^*$  is usually nullified by the extra computational effort needed by Newton's method to calculate the Hessian at each iteration.

The performance of the conjugate gradient and quasi-Newton methods are comparable; both require approximately the same number of iterations (16 and 20, respectively), objective function calculations (214 and 151, respectively) and first-order sensitivity calculations (35 and 41, respectively). This suggests that either method would



be a suitable choice for a minimization technique in cases where the second-order objective function sensitivities are not readily available.

As shown in Fig. A.7, Newton's method has the greatest rate of convergence of the four methods throughout the entire minimization process. In the vicinity of  $\Phi^*$ , however, the quasi-Newton and conjugate gradient methods attain the same convergence rate as Newton's method; in the case of the quasi-Newton method, this is indicative of an accurate Hessian approximation. The steepest descent method has the slowest convergence rates; while the convergence rates of the Newton, quasi-Newton, and conjugate gradient methods are all superlinear, the steepest descent convergence rate is linear, where as  $k$  becomes large,

$$\frac{|F(\Phi^{k+1}) - F(\Phi^*)|}{|F(\Phi^k) - F(\Phi^*)|} = C. \quad (\text{A.39})$$

This slow convergence rate makes the steepest descent method an unsuitable minimization technique for most nonlinear problems.

Minimization Method	Total Iterations	$F(\Phi)$ Evaluations	$\partial F(\Phi)/\partial \Phi_p$ Evaluations	$\partial^2 F(\Phi)/\partial \Phi_p \partial \Phi_q$ Evaluations
Newton	9	119	19	10
Quasi-Newton	16	214	35	0
Steepest Descent	926	7013	927	0
Conjugate Gradient	20	151	41	0

Table A.1: Performance of Newton, quasi-Newton, steepest descent and conjugate gradient minimization techniques applied to minimize Rosenbrock's function.

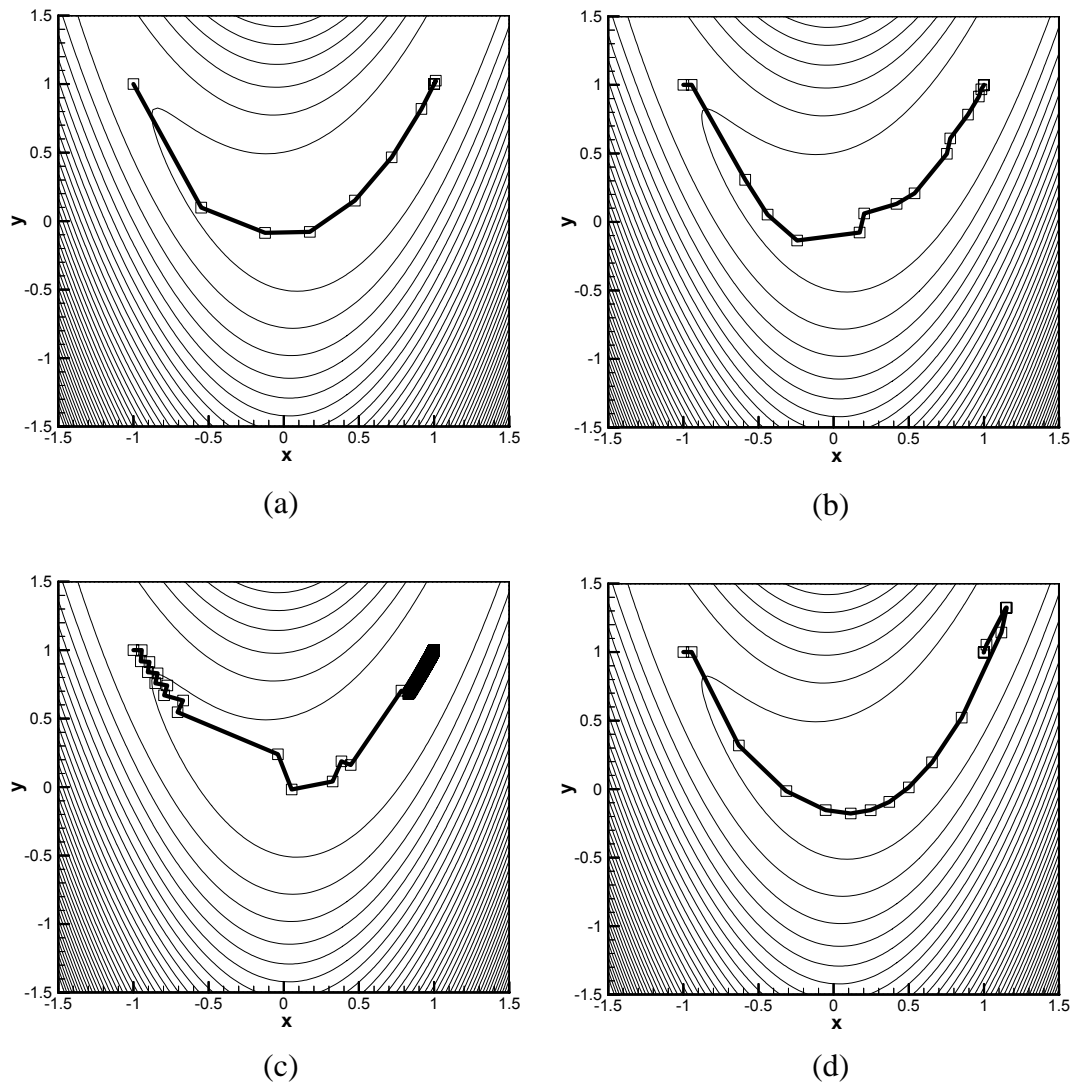


Figure A.6: Minimization paths for different gradient-based optimization techniques:  
(a) Newton's method, (b) Quasi-Newton's method, (c), Steepest-Descent, and (d)  
Conjugate Gradient Method.

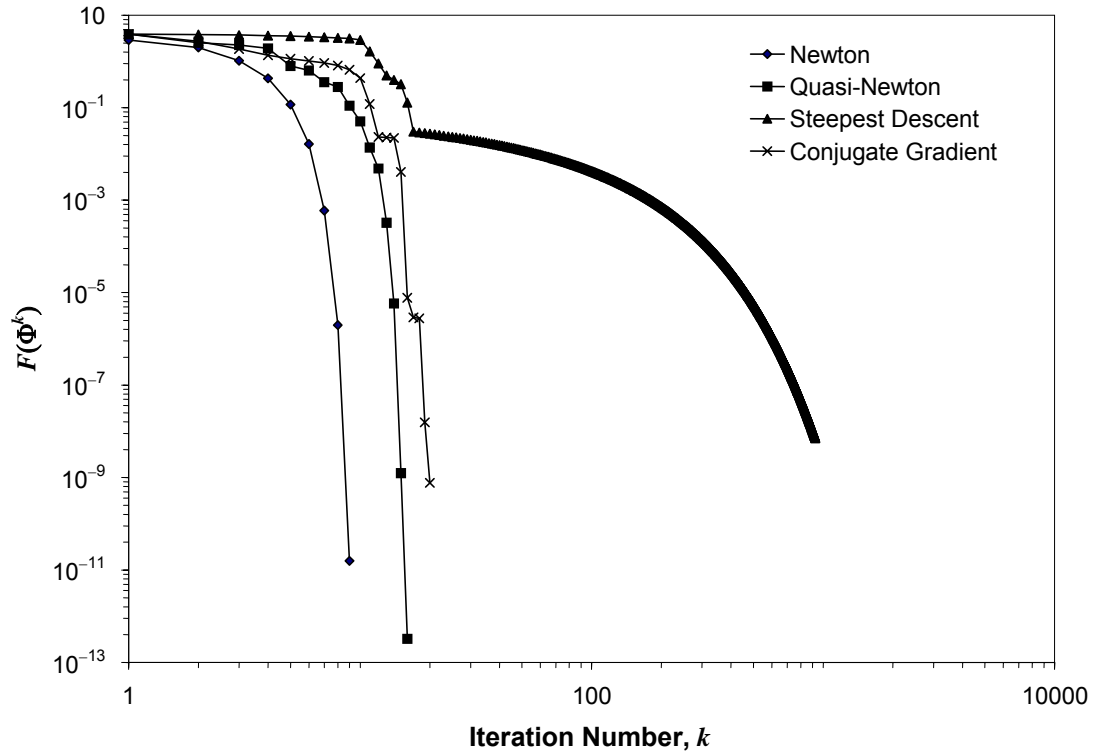


Figure A.7: Plot of  $F(\Phi)$  with respect to iteration number for Newton, quasi-Newton, steepest descent and conjugate gradient minimization techniques.

### A.3 Constrained Minimization Techniques

It is often necessary to restrict the values of  $\Phi$  to lie within some fixed domain (called a “feasible region”) by specifying constraints. In the context of radiant enclosure design, constraints are used to prevent heater outputs from assuming negative values, to maintain an unobstructed enclosure geometry (so that any two points on the enclosure surface can “see” each other), and to restrict the enclosure dimensions to be less than some maximum size. In this application, all the constraints are linear inequalities, so the minimization problem can be written as

$$\text{Minimize } F(\Phi), \quad \Phi \in \mathcal{R}^n, \quad (\text{A.40})$$

subject to

$$c_i(\Phi) \geq 0, \quad i = 1 \dots m. \quad (\text{A.41})$$

The  $i^{\text{th}}$  inequality constraint is said to be *active* at  $\Phi^k$  if  $c_i(\Phi^k) = 0$ ; otherwise, the constraint is *inactive*.

It is often useful to incorporate the objective function and constraints into a single function called the Lagrangian function,

$$L(\Phi, \xi) = F(\Phi) + \xi^T c(\Phi). \quad (\text{A.42})$$

Vectors  $\xi$  and  $c(\Phi)$  contain the Lagrangian multipliers and inequality constraints, respectively, with the  $i^{\text{th}}$  Lagrangian multiplier,  $\xi_i$ , corresponding to the  $i^{\text{th}}$  constraint,  $c_i(\Phi)$ . It can be shown that the  $n$ -dimensional constrained minimization problem defined by Eqs. (A.40) and (A.41) is equivalent to the  $n + m$  dimensional unconstrained problem where the objective is to find the vectors  $\Phi^*$  and  $\xi^*$  such that

$$L(\Phi^*, \xi^*) = \text{Min}[L(\Phi, \xi)], \quad \Phi \in \mathcal{R}^n, \quad \xi \in \mathcal{R}^m. \quad (\text{A.43})$$

A solution to Eq. (A.43) thus satisfies the first-order necessary conditions for a local minimum,

$$\frac{\partial L(\Phi^*, \xi^*)}{\partial \Phi_p} = \frac{\partial F(\Phi^*)}{\partial \Phi_p} + \sum_{i=1}^m \xi_i \frac{\partial c_i(\Phi^*)}{\partial \Phi_p} = 0, \quad p = 1 \dots n. \quad (\text{A.44})$$

Assuming the inequalities are defined in a form consistent with Eq. (A.41), the Lagrangian multipliers are strictly non-positive. If the  $i^{\text{th}}$  constraint is active, then  $\xi_i$  is equal to some negative number, and  $\xi_i = 0$  otherwise. Lagrangian multipliers can also be interpreted as marginal values, and are a linear approximation of how relaxing the active constraints would affect the minimization problem. In particular,  $F(\Phi)$  would decrease (improve) by approximately  $\xi_i$  if  $c_i(\Phi^*)$  was allowed to increase by one unit.

This can be most clearly seen if the inequality constraints are simple bounds of the form

$$c_{2i}(\Phi) = \Phi_i - l_{bi} \geq 0, \quad i = 1 \dots n, \quad (\text{A.45})$$

and

$$c_{2i+1}(\Phi) = u_{bi} - \Phi_i \geq 0, \quad i = 1 \dots n, \quad (\text{A.46})$$

resulting in  $m = 2n$  inequality constraints. If  $\Phi_p$  lies in between  $l_{bp}$  and  $u_{bp}$  then  $\partial F(\Phi^*)/\partial \Phi_p = 0$ , and in order for Eq. (A.44) to hold true,  $\xi_p = 0$ . On the other hand, if  $\Phi_p$  is at a lower bound, then  $\partial c_{2p}(\Phi^*)/\partial \Phi_p = 1$ ,  $\partial F(\Phi^*)/\partial \Phi_p > 0$  and from Eq. (A.44),  $\xi_{2p} = -\partial F(\Phi^*)/\partial \Phi_p$ . Finally, if  $\Phi_p$  is at an upper bound then  $\partial c_{2p+1}(\Phi^*)/\partial \Phi_p = -1$ ,  $\partial F(\Phi^*)/\partial \Phi_p < 0$  and  $\xi_{2p+1} = \partial F(\Phi^*)/\partial \Phi_p$ . These three scenarios are shown in Fig. A.8, for the objective function  $F(x, y) = -0.25y^2(y - 2)(y + 2) + x^2$  constrained by

$$c_1(y) = 3 - y \geq 0, \quad (\text{A.47})$$

and

$$c_2(y) = y + 3 \geq 0. \quad (\text{A.48})$$

At the local minimum  $F(0, 0) = 0$ ,  $\nabla F(0, 0) = \{0, 0\}^T$ , so all the Lagrange multipliers equal zero, i.e.  $\xi_i = 0, i = 1 \dots 2n$ . At a second local minimum,  $F(0, 3) = -11.25$ , the upper constraint is active so  $\xi_1 = \partial f(0, 3)/\partial y = -21$  and  $\xi_2 = 0$ . In other words, if the upper constraint was relaxed by one unit, this local minimum could be further reduced by approximately 21 units from  $-11.25$  to  $-32.25$ . (In fact, changing the first constraint to  $4 - y \geq 0$  results in a new local minimum  $F(0, -4) = -48$ .) Finally, at the third local minimum,  $F(0, -3) = -11.25$ ; since the lower constraint is active at this local minimum,  $\xi_1 = 0$  and  $\xi_2 = -\partial f(0, -3)/\partial y = -21$ .

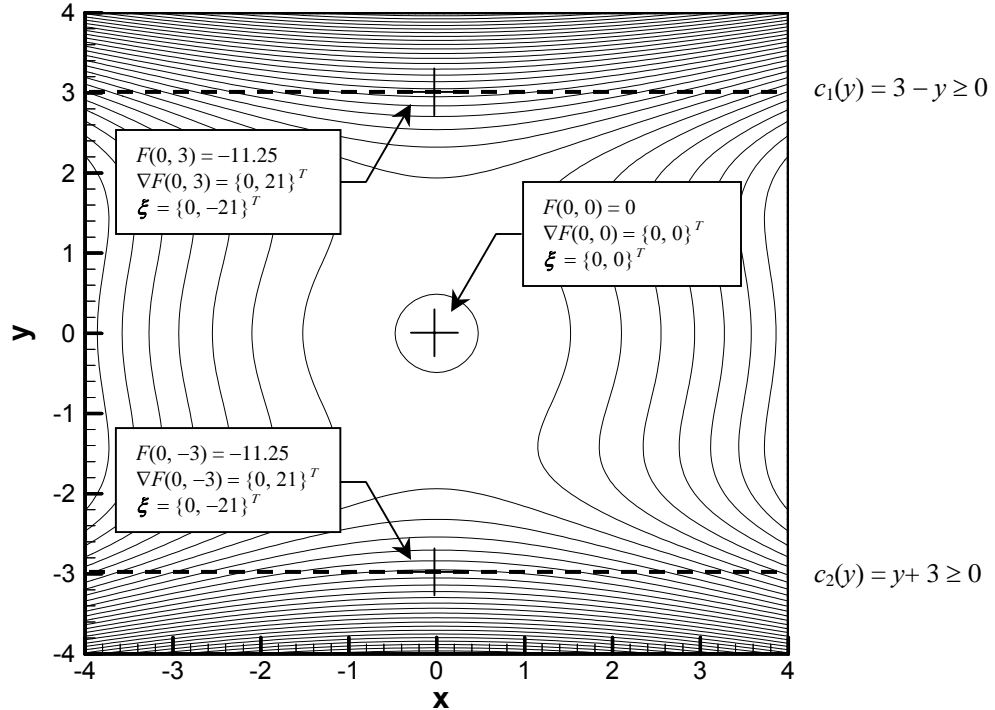


Figure A.8: Lagrange multipliers for constrained minimization problems having simple bounds, plot of  $F(x, y) = -0.25 y^2(y-2)(y+2) + x^2$ .

Constraints are most often enforced by modifying the unconstrained minimization techniques discussed in the previous section. Two such techniques, gradient projection and active set methods, are discussed in the following sections.

### A.3.2 Gradient Projection Method

The gradient projection method is one of the simplest techniques used to enforce inequality constraints, and can be easily integrated into the unconstrained minimization techniques discussed in the previous section.

The method works as follows; at each iteration, the prospective step  $\Phi^k + p^k \alpha^k$  is checked to see if it lies within the feasible region. If not, then a point  $\bar{\Phi}^k$  is found by projecting  $\Phi^k + p^k \alpha^k$  onto the boundary of the feasible region in such a way that  $\|\bar{\Phi}^k - (\Phi^k + p^k \alpha^k)\|_2$  is minimized. This new point is then used to define a feasible direction,  $d^k = \bar{\Phi}^k - \Phi^k$ , and finally  $\Phi^{k+1}$  is found by taking a step in the  $d^k$  direction with a step length  $s^k$ , so  $\Phi^{k+1} = \Phi^k + d^k s^k$ . Most often,  $s^k$  equals unity, and  $\Phi^{k+1}$  is then the projected point,  $\bar{\Phi}^k$ .

To demonstrate the gradient projection method, consider minimization of the quadratic function of  $F(x, y) = (x - 2y)^2 + 2x^4$  subject to  $y \geq 2 - x/4$  and  $y \geq 5x - 10$ , as shown in Fig. A.9. The minimization is carried out using steepest descent coupled with an exact line search. Minimization starts at  $\Phi^0 = \{3, 3\}^T$ , and the first iteration identifies a new point,  $\Phi^1$ , that satisfies the constraints. The next point found with the steepest descent direction, however, violates the first constraint and is therefore said to be

*infeasible*. Instead, an alternate feasible point,  $\bar{\Phi}^1$ , is identified by projecting  $\Phi^1 + p^1 \alpha^1$  onto  $y = 2 - x/4$ , which also defines a new feasible search direction,  $d^1 = \bar{\Phi}^1 - \Phi^1$ . In this example, the step length  $s^1$  is set equal to unity, and  $\Phi^2 = \Phi^1 + d^1 s^1 = \bar{\Phi}^1$ . This procedure is repeated until a local minimum is found.

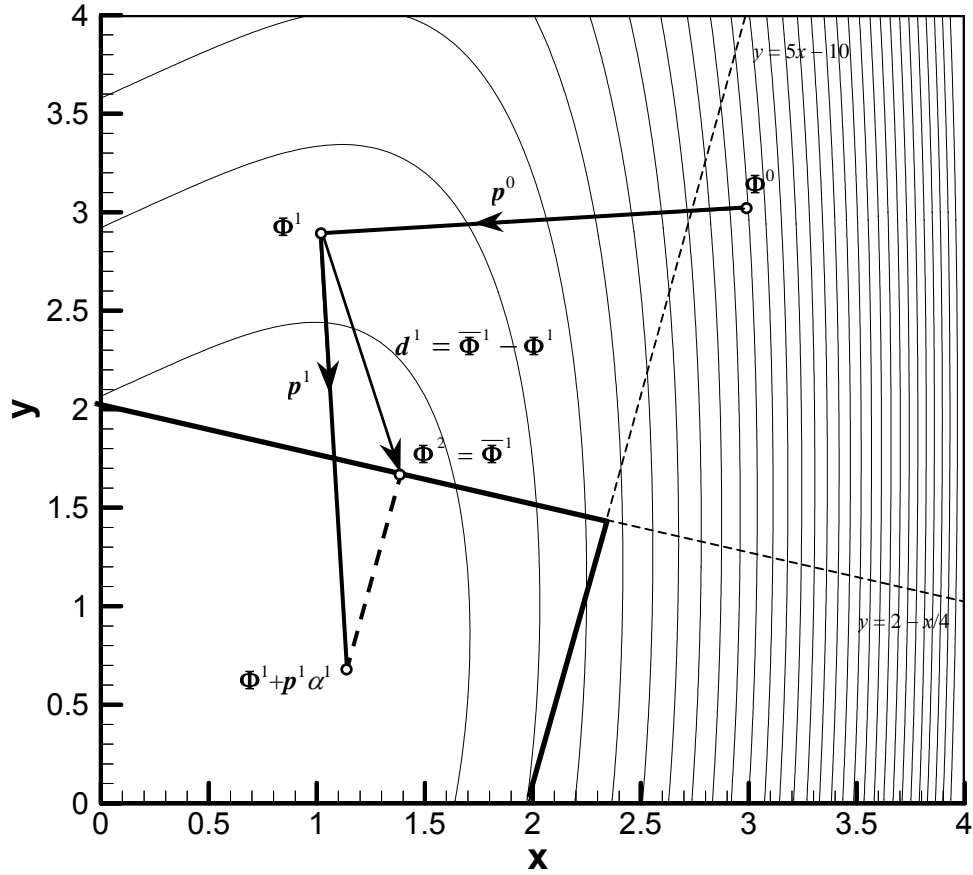


Figure A.9: First two steps of the minimization of  $F(x, y) = (x - 2y)^2 + 2x^4$  subject to  $y > 2 - x/4$  and  $y > 5x - 10$  by steepest descent using an exact line search, coupled with the gradient projection method.



### A.3.2 Active Set Method

Active set methods are derived based on the observations that optimization problems containing only equality constraints are generally easier to solve than problems where some or all of the constraints are inequalities; furthermore, if it were known which inequality constraints were active at  $\Phi^*$  beforehand (i.e. which Lagrangian multipliers are nonzero), then the optimization problem could be treated as an equality-constrained problem. In the special case where the constraints are simple bounds the variables corresponding to the  $t$  active constraints could be set equal to their active bounds ( $\Phi_i = l_{bi}$ , for example), and the  $n$ -dimensional constrained optimization problem would then be reduced to an  $(n - t)$ -dimensional unconstrained optimization problem.

Since the active set of constraints at  $\Phi^*$  is generally unknown, however, active set methods work by first guessing a “working set” of constraints, which is updated after each successive iteration until the correct working set has been identified and the problem has been solved. These methods work according to the following scheme. Suppose that at the  $k^{th}$  iteration,  $t$  variables are fixed at their constraints:

1. The “reduced” gradient vector and Hessian matrix (or approximate Hessian matrix) corresponding to the  $n - t$  unconstrained minimization problem are formed, and a search direction,  $\mathbf{p}^k$ , is found using either Eq. (A.11) or Eq. (A.12).
2. The closest constraint in the  $\mathbf{p}^k$  direction is identified and the smallest step size that makes  $\Phi^k + \alpha^k \mathbf{p}^k$  infeasible,  $\alpha^{k_{crit}}$ , is calculated.

3. A step size,  $\alpha^k$ , is chosen using one of the methods previously described. If  $\alpha^k > \alpha^{krit}$ , the minimization has “run into” the critical constraint, so  $\alpha^k$  is set equal to  $\alpha^{krit}$  and the constraint is added to the active set.
4. An estimate for  $\xi^*$ ,  $\xi^k$ , is formed using  $\nabla F(\Phi^k)$  and, if available,  $\nabla^2 F(\Phi^k)$ .
5. The active set is updated, and if  $\xi_i^k \geq 0$  then the corresponding constraint is removed from the active set.

This procedure continues until  $\Phi^k = \Phi^*$  and  $\xi^k = \xi^*$ .

Using the estimated Lagrangian multipliers to modify the active set in step 5 ensures that every search direction calculated in Step 2 is a descent direction, providing that the same order of method used to predict  $\mathbf{p}^k$  is also used to predict  $\xi^k$ . For example, suppose the constraints are all simple bounds of the form  $l_{bi} \leq \Phi_i$  and  $\Phi_i \leq u_{bi}$ . (This is not consistent with the form of Eq. (A.41), but is more convenient in the present context.) If a quasi-Newton routine is used to calculate  $\mathbf{p}^k$ , then the Lagrangian multipliers corresponding to constraints in the active set are estimated by  $\xi_i^k = \partial F(\Phi^k) / \partial \Phi_p$  if  $\partial F(\Phi^k) / \partial \Phi_p < 0$ , where  $\Phi_p$  is constrained by  $c_i$ . If one of the Lagrangian multipliers were positive, then the objective function could be reduced by moving further into the feasible region in a direction away from the bound, so the corresponding constraint is removed from the active set.

Once  $\xi^k$  is calculated, the reduced gradient and reduced approximate Hessian matrix are formed by removing all elements corresponding to the  $t$  variables that are fixed at their bounds. The search direction found by solving the resulting  $n - t$  system of linear

equations is in a direction parallel to the active constraints; in this way, the minimization method moves along the active constraints until they are either removed from the active set (because the corresponding Lagrange multiplier estimates are nonnegative) or until  $\Phi^*$  and  $\xi^*$  are found.

The active set method is demonstrated by applying it to minimize Rosenbrock's function,  $F(x, y) = (x - 1)^2 + 10(x^2 - y)^2$ , subject to  $-1.5 \leq x \leq 0.5$  and  $0 \leq y \leq 1.5$  as shown in Fig. A.10. Minimization is carried out using the quasi-Newton method with a unit step size, and starts from  $\Phi^0 = \{-1.5, 1.5\}^T$ . The constraint  $y \leq 1.5$  is immediately added to the working set, and is only removed once the  $y$ -component of the gradient becomes positive, meaning that the  $F(\Phi)$  could be reduced by moving away from  $y = 1.5$  in a feasible direction. Next, the minimization path intersects the constraint  $y \geq 0$  several times, which is not added to the working set since  $F(\Phi)$  can be reduced each time by moving in the positive  $y$ -direction. Finally, the constraint  $x \leq 0.5$  is encountered; since the objective function could be improved by relaxing this constraint, it is added to the working set, and the remaining few steps are equivalent to performing a single-variable minimization on  $F(1.5, y)$  with respect to  $y$ . The minimum value of  $F(x, y)$  is equal to 0.25, with  $\Phi^* = \{0.5, 0.25\}^T$ .

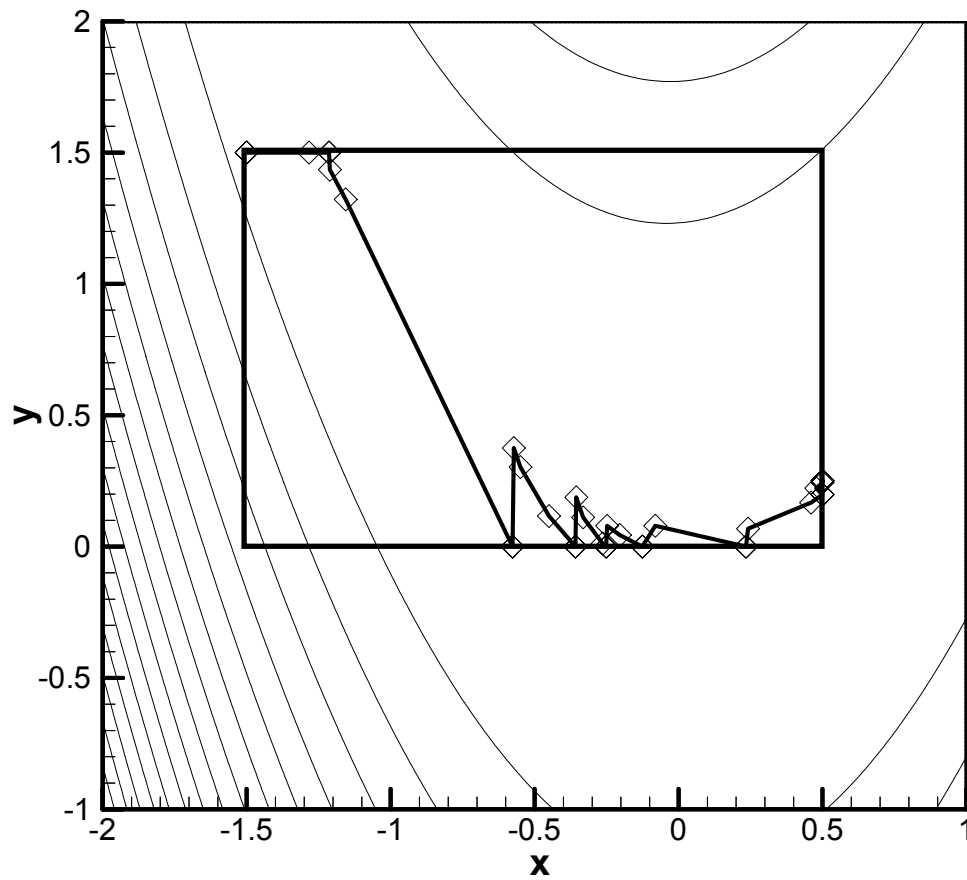


Figure A.10: Minimization of Rosenbrock's function,  $F(x, y) = (x - 1)^2 + 10(x^2 - y)^2$ , subject to  $-1.5 \leq x \leq 0.5$  and  $0 \leq y \leq 1.5$ , using the quasi-Newton method with a unit step size coupled with an active set method.

#### A.4 Solution of Multimodal Problems

If the objective region is strictly convex over the feasible region, i.e.

$$F[s\Phi_1 + (1-s)\Phi_2] < sF(\Phi_1) + (1-s)F(\Phi_2), \quad 0 \leq s \leq 1, \quad \Phi_1 \neq \Phi_2, \quad (\text{A.49})$$

and if the feasible region is a convex set, it can be shown that any local minimum that lies within the feasible region is a global minimum.

In this application, the constraints are all assumed to be linear inequalities so the feasible region is always a convex set. The shape of the objective function, on the other hand, is most often unknown and therefore it cannot be shown to be convex over the feasible region. Accordingly, a stationary point identified by the minimization techniques in the above sections is not necessarily a global minimum. For example, consider optimization of the objective function

$$F(x, y) = (2x - 10x^3 - 10y^5)e^{-x^2-y^2} - \frac{1}{3}e^{2x-1-x^2-y^2} \quad (\text{A.50})$$

over the feasible region defined by

$$y \geq 1 - x, \quad (\text{A.51})$$

$$y \leq 5 - 5/2x, \quad (\text{A.52})$$

$$y \leq 2 + 1/2x, \quad (\text{A.53})$$

$$y \geq 0, \quad (\text{A.54})$$

and

$$x \geq 0. \quad (\text{A.55})$$

As shown in Fig. A.11, the global minimum is located at  $\Phi^{*1} = \{0, 1.5802\}^T$  with  $F(\Phi^{*1}) = -8.1218$ , with two other local minima at  $\Phi^{*2} = \{1.2696, 0\}^T$  with  $F(\Phi^{*2}) = -3.8862$  and  $\Phi^{*3} = \{1.8932, 0.4020\}^T$  with  $F(\Phi^{*3}) = -1.8343$ , and a local maximum at  $\Phi^{*4} =$

$\{1.3070, 1.7326\}^T$  with  $F(\Phi^{*4}) = -1.5984$ . A single minimization carried out using any of the methods discussed in the previous sections will find only one of the stationary points (possibly the local maximum) depending on the search method and starting point. If the form of  $F(\Phi)$  were to be unknown, it would be impossible to know how many minima are contained in the feasible region, or if a particular local minimum was the global minimum, without plotting  $F(\Phi)$  over the feasible region. Furthermore, this procedure is often computationally expensive and very difficult to do in problems where  $n > 2$ .

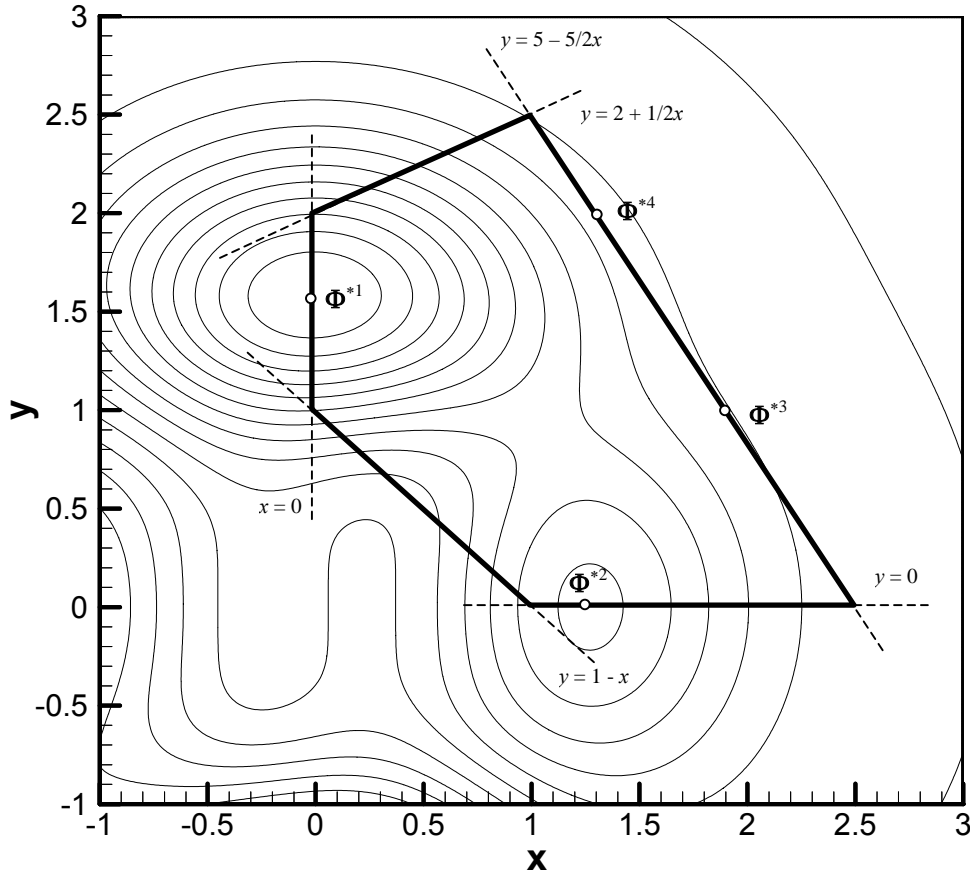


Figure A.11: Objective function of Eq. (A.50) and feasible region defined by constraints in Eqs. (A.51) – (A.55).

In the context of gradient-based minimization, the most common way of solving this problem is employ a multi-start algorithm, which searches for the global minimum by performing multiple local minimizations from different initial points. Although this method does not guarantee a global minimum (the probability of finding the global minimum becomes unity as the number of restarts approaches infinity), it does increase the chances of finding a good solution.

The simplest type of multi-start algorithm uses a set of initial points that are uniformly distributed throughout the feasible region. This procedure is highly inefficient, however, because many of the minimizations will converge to the same local minimum. A more efficient multi-start routine has recently been contributed by Urgay et al. (2002).

**Appendix B:**  
**Solution of Inverse Radiant Enclosure Design**  
**Problems through Regularization**



## B.1 Introduction

The majority of radiant enclosure design problems discussed in the literature arise from applications where heaters located within radiant enclosures are used to impose a temperature and heat flux distribution over a design surface. In many industrial processes, for example, radiant enclosures are used to carry out the heat treatment of products located on the design surface. When designing these enclosures it is necessary to calculate the heater settings that produce the required heat flux and temperature distribution over the products throughout the process. Radiant enclosures are also used to simulate the thermal response of objects within fires; for example, a nuclear waste storage vessel immersed within a pool fire can be modeled by placing the vessel within a radiant enclosure. In order to carry out these experiments, heat flux distributions over the heater surface must be found that produce conditions over the design surface that simulate conditions measured experimentally within a fire environment.

Traditionally, these design problems have been solved in their *implicit* form, where one boundary condition is specified over the design surface, and the designer repeatedly adjusts the enclosure configuration (relying only on his or her intuition and experience) until the distribution of the unspecified boundary condition matches the desired distribution within an allowable tolerance. This trial-and-error approach is very time-consuming, and the final solution quality is usually limited.

More recently, inverse design methodologies have been applied to overcome these drawbacks. In this approach, the inverse design problem is solved in its *explicit* form, where both the desired heat flux and temperature distributions are enforced over the

design surface while the heat flux and temperature distributions over the heater surface are not specified. Written this way, inverse design problems are *ill-posed* because a unique solution to the design problem may not exist; instead, there could be either multiple solutions (in this application, multiple heat flux distributions over the heater surface that produce the desired boundary conditions over the design surface), or no solution at all. Discretizing the equations that govern the physical system produces a system of linear equations that in general contain an unequal amount of linearly-independent equations and unknowns. Systems with more unknowns than equations have an infinite number of solutions, while those with more equations than unknowns are unlikely to have a non-trivial solution that satisfies the equations exactly. Accordingly, traditional linear algebra techniques, such as LU decomposition or Gauss-Seidel iteration, cannot be used to identify a useful solution to the design problem.

Instead, regularization techniques are employed to identify the most useful solutions from the infinite set that satisfy these equations. (The “usefulness” of a solution is often quantified by the  $L_2$  norm of the solution vector, with the most useful solutions having the smallest solution norms.) The useful solutions are identified by solving a sequence of well-posed problems that are related to the ill-posed inverse design problem. Unlike the original ill-posed problem, each of these well-posed problems has a unique solution; however, since the well-posed problems differ from the original ill-posed problem, substituting the solution of one of these well-posed problems back into the original ill-conditioned set of linear equations results in a non-zero residual vector. The well-posed problems that are most closely related to the original ill-posed problem

characteristically have solution vectors with relatively large norms, but produce small residuals when substituted back into the original ill-conditioned set of equations. On the other hand, well-posed problems formed by extensively modifying the ill-posed problem have comparatively small solution norms, but produce large residual norms. Thus, selecting one of these solutions represents a compromise between the solution regularity and solution accuracy, which are indicated by the norms and the solution and residual vectors, respectively.

In this appendix, the definition of an ill-posed problem is presented and the source of the ill-conditioned system of linear equations is described. The three regularization methods that are commonly used to solve inverse radiant enclosure design problems, truncated singular value decomposition (TSVD), Tikhonov regularization, and iterative conjugate gradient regularization, are then described in detail. Finally, these methods are demonstrated by using them to solve an example radiant enclosure design problem.

## B.2 Ill-Posed Problems

The study of ill-posed problems began with the distinction made by Hadamard (1923) between well-posed and ill-posed mathematical problems. In particular, he defined the three criteria that must be satisfied for a mathematical problem to be well-posed:

1. The problem must have a solution,
2. This solution must be unique, and
3. The solution must be stable under small changes to the input data.

Any problem that does not satisfy all of the above criteria is said to be ill-posed. Ill-posed problems were initially thought to be artificial in the sense that they did not describe natural systems, and furthermore solutions to these problems would be physically meaningless. It has been shown that this is not the case and that, in fact, many of the problems that arise in diverse fields of natural science and engineering are fundamentally ill-posed. In particular, most inverse engineering design problems are ill-posed because there may exist many possible solutions of varying quality, or no satisfactory solution at all.

For example, consider the inverse radiant enclosure design problem described in Chapter 2, where the objective is to find a heat flux distribution over the heater surface that produces the desired temperature and heat flux distributions over the design surface. When the inverse problem is written in its implicit form, it was shown that the radiosity distribution is governed by a Fredholm integral equation of the second-kind, which has the form

$$f(s) = g(s) + \int_a^b f(s') k(s, s') ds', \quad (\text{B.1})$$

where both  $g(s)$  and the kernel,  $k(s, s')$  are known functions while  $f(s)$  is an unknown function. In most cases there exists only one unique solution for  $f(s)$  that satisfies Eq. (B.1), and accordingly these problems are usually well-posed.

If the inverse problem is written explicitly, on the other hand, the radiosity distribution is governed in part by a Fredholm integral equation of the first-kind having the general form

$$g(s) = \int_a^b f(s') k(s, s') ds', \quad (\text{B.2})$$

where again  $g(s)$  and  $k(s, s')$  are known while  $f(s')$  is unknown. Clearly, an infinite number of different functions could be substituted for  $f(s')$  to satisfy Eq. (B.2); identifying those that provide useful solutions to the inverse problem, however, is usually far more challenging than solving a related well-posed problem.

Analytical solutions of integral equations are usually intractable and therefore solutions must be found using numerical techniques. Discretizing the domain of either one of these integral equations produces a system of linear equations that can be rewritten as a matrix equation,

$$\mathbf{A} \mathbf{x} = \mathbf{b}. \quad (\text{B.3})$$

If the governing equations are well-posed,  $\mathbf{A}$  is most often well-conditioned. In this case, the number of independent equations equals the number of unknowns, so Eq. (B.3) can be solved using traditional linear algebra techniques such as LU decomposition and Gauss-Seidel iteration.

On the other hand, if the governing equations are ill-posed, then the solution of Eq. (B.3) is far more difficult to carry out. If there is one solution that exactly satisfies

the governing equations and also a set of solutions that almost satisfy the governing equations with a small residual, then  $\mathbf{A}$  is said to be ill-conditioned. In this situation, several of the equations that comprise the linear system are very similar, and  $\mathbf{A}$  is very nearly rank-deficient. If there are multiple solutions to the governing equations, as is often the case in inverse radiant enclosure problems, then there are more unknowns than linearly-independent equations (i.e. more columns than rows) and  $\mathbf{A}$  is singular. In particular, if the number of unknowns exceeds the number of equations by  $k$ , then  $\mathbf{A}$  is said to have a *nullity* of  $k$ . In both of these two cases, special regularization methods must be employed to solve Eq. (B.3).

### B.3 Regularization Methods

Many different regularization methods have been developed to solve ill-conditioned sets of linear equations; Hansen (1998) and Vogel (2002) summarize these methods, while França et al. (2002) focuses on how regularization is used to solve inverse radiant enclosure design problems. The three regularization methods most commonly used to solve these types of problems are truncated singular value decomposition (TSVD), Tikhonov regularization, and iterative conjugate gradient regularization. Although each method solves the inverse problem in a fundamentally different way, all of them include a heuristic parameter that controls the amount of regularization used to find the solution. As a result, a set of solutions with varying degrees of regularity and accuracy are generated, and it is left to the designer to select the one that best satisfies the requirements of the inverse design problem.

#### B.3.1 Truncated Singular Value Decomposition

The TSVD method is based on the singular value decomposition of the  $\mathbf{A}$  matrix,

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T, \quad (\text{B.4})$$

where  $\mathbf{U}$  is an orthogonal matrix with  $m$  rows and  $m + k$  columns,  $\mathbf{W}$  is a diagonal matrix with  $m + k$  rows and columns, and  $\mathbf{V}$  is the transpose of an  $(m + k) \times (m + k)$  orthogonal matrix. The diagonal elements of  $\mathbf{W}$ ,  $w_i$ , are the singular values, and are strictly non-negative. The degree of difficulty associated with solving a linear system of equations can be anticipated by examining the singular values, and specifically the *condition number*,

$$Cond(\mathbf{A}) = \frac{w_{\max}}{w_{\min}}, \quad (\text{B.5})$$

where  $w_{\max}$  and  $w_{\min}$  are the largest and smallest singular values, respectively. If all the singular values are of approximately the same magnitude and the condition number is relatively small, then  $\mathbf{A}$  is said to be *well-conditioned*, and the matrix equation can be solved by back-substitution,

$$x_s = \sum_{i=1}^{m+k} \left[ \frac{V_{si}}{w_i} \sum_{j=1}^m U_{ji} b_j \right]. \quad (\text{B.6})$$

On the other hand, if some of the singular values are very small relative to others, the condition number will be large and  $\mathbf{A}$  is said to be *ill-conditioned*; in inverse radiant enclosure design problems, these small singular values arise from the existence of multiple solutions that, when substituted into the forward problem, produce heat flux and temperature distributions over the design surface that closely match the desired distributions. In the case where the number of unknowns exceeds the number of equations by  $k$ , at least  $k$  of the singular values equal zero; the condition number is then infinite and  $\mathbf{A}$  is said to be *singular*. (Assuming the remaining  $m$  equations are linearly independent,  $k$  is said to be the *nullity* of  $\mathbf{A}$ .) Attempting to solve the system of equations using Eq. (2.18) would result in either a solution having large oscillations if  $\mathbf{A}$  were ill-conditioned, or division by zero if  $\mathbf{A}$  were singular.

In *TSVD*, the singular values that are less than some user-defined criterion are negated (or truncated) by setting the corresponding  $1/w_i$  terms in Eq. (B.6) equal to zero, which creates a related well-posed problem that is related to the original ill-posed problem. If few singular values are truncated, the integrity of the original problem is



maintained; the norm of the residual vector,  $\|\delta\|_2$  tends to be small while the radiosity and heat flux distributions usually have very irregular distributions, and consequently the norm of the solution vector,  $\|\mathbf{x}\|_2$ , is quite large. Truncating more of the singular values acts to regularize the solution, which lowers  $\|\mathbf{x}\|_2$  at the cost of increasing  $\|\delta\|_2$ . It is then left to the designer to find a solution that is both smooth enough to be practically implemented in a design setting and sufficiently accurate so that the original equations governing the radiation heat transfer within the enclosure are obeyed.

### B.3.2 Tikhonov Regularization

This regularization technique was first proposed by Tikhonov (1975) for solving inverse conduction problems; it makes use of *a priori* knowledge of the approximate solution size and smoothness to form a sequence of well-posed problems that are related to the original ill-posed problem. By adjusting the amount of regularization, the analyst first generates a set of potential solutions, and then selects the one that offers the most acceptable compromise between regularity and accuracy.

The method uses optimization principles to solve the inverse problem, and is based on the observation that the objective of the regularization process is to find a solution that is both accurate and regular. Both of these goals could be achieved independently by minimizing two separate objective functions. The most accurate solution is found by minimizing the magnitude of the residual vector,

$$F_a(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2. \quad (\text{B.7})$$

If the linear system is rank-deficient, however,  $F_a(\mathbf{x})$  will not have a unique local minimum; instead, there are an infinite set of solutions,  $\{\mathbf{x}_a^*\}$  that minimize  $F_a(\mathbf{x})$ . In order to visualize such a situation, consider the rank-deficient set of linear equations given by,

$$\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -4 \\ -2 \end{bmatrix}. \quad (\text{B.8})$$

As shown in Fig. B.1, any solution located along the line  $2x_1 + x_2 = -2$  minimizes  $F_a(\mathbf{x})$ .

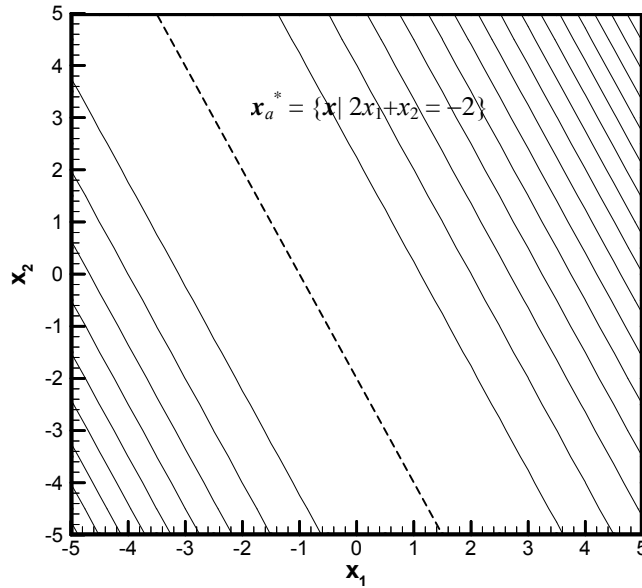


Figure B.1: Plot of  $F_a(\mathbf{x})$  for the matrix equation defined in Eq. (B.8)

Without considering solution accuracy, the most regular solution is found by minimizing the objective function

$$F_b(\mathbf{x}) = \sum_{i=0}^{\gamma} \rho_i \left\| \mathbf{L}^i (\mathbf{x} - \mathbf{x}^0) \right\|_2^2, \quad (\text{B.9})$$

where  $\mathbf{x}^0$  is an *a priori* solution guess,  $\gamma$  is the order of curve smoothness,  $\rho_i$  is a heuristic weight, and  $\mathbf{L}^i$  approximates the  $i^{\text{th}}$  derivative operator. The order of curve smoothness is

chosen based on the estimate of the solution shape; for example, if the distribution of  $\mathbf{x}$  is expected to be uniform, then  $\gamma = 0$ , and  $\mathbf{L}^0$  is the identity matrix. On the other hand, if a purely quadratic solution shape is expected, then  $\gamma = 2$ ,  $\rho_0 = \rho_1 = 0$ , and a suitable choice for  $\mathbf{L}^2$  is the tri-diagonal matrix

$$\mathbf{L}^2 = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix}. \quad (\text{B.10})$$

In this instance, the solution that minimizes  $F_b(\mathbf{x})$ ,  $\mathbf{x}_b^*$ , would be a smooth quadratic function. Unlike the objective function defined in Eq. (B.7),  $F_b(\mathbf{x})$  has one unique global minimum, as shown in Fig. B.2 for the system of linear equations defined in Eq. (B.8) with  $\gamma = 0$  and  $\mathbf{x}_0 = \{0, 0\}^T$ .

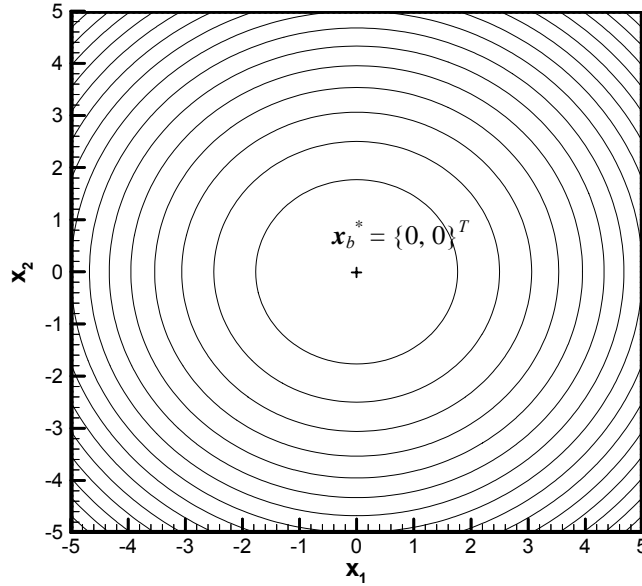


Figure B.2: Plot of  $F_b(\mathbf{x})$  for the matrix equation defined in Eq. (B.8).

Unfortunately, the smoothest solution and the most accurate solution are different since the global minimum of  $F_b(\mathbf{x})$  is generally not in the set of solutions that minimize  $F_a(\mathbf{x})$ . Instead, a compromise between solution accuracy and regularity is found by minimizing a new objective function defined by

$$F(\mathbf{x}) = F_a(\mathbf{x}) + cF_b(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \sum_{i=0}^{\gamma} \beta_i \|\mathbf{L}^i(\mathbf{x} - \mathbf{x}^0)\|_2^2, \quad (\text{B.11})$$

where  $\beta_i = c\rho_i$  is the heuristic regularization parameter.

Moreover, since adding an objective function with an infinite set of local minima to an objective function having a unique global minimum produces another objective function with a unique global minimum, Eq. (B.11) transforms the original ill-posed problem with an infinite set of solutions into a well-posed problem having a unique solution. This is demonstrated by comparing the plots of  $F_a(\mathbf{x})$  and  $F(\mathbf{x})$  shown in Figs. B.1 and B.3, respectively, for the system of linear equations defined in Eq. (B.8).

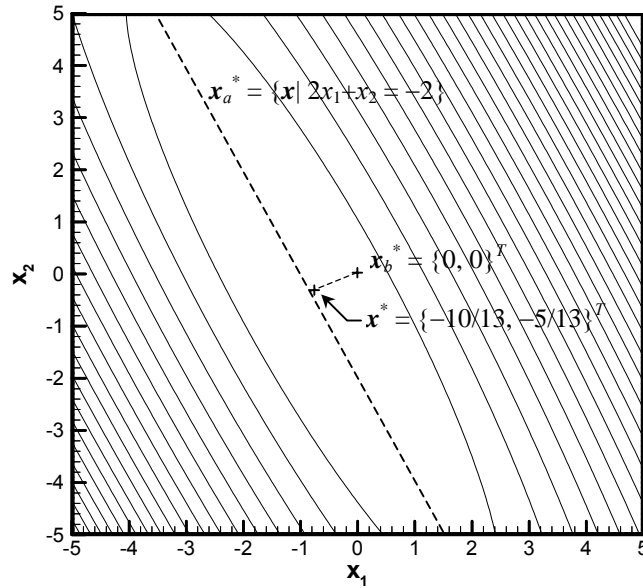


Figure B.3: Plot of  $F_a(\mathbf{x}) + \beta F_b(\mathbf{x})$ ,  $\beta = 1$  for the matrix equation defined in Eq. (B.8).

Since  $F(\mathbf{x})$  is a quadratic, the global minimum can be found analytically by satisfying the first order necessity conditions at  $\mathbf{x}^*$ ,

$$\nabla F(\mathbf{x}) = 2\mathbf{A}^T (\mathbf{A}\mathbf{x}^* - \mathbf{b}) + 2\sum_{i=0}^{\gamma} \beta_i \mathbf{L}^i \mathbf{L}^T (\mathbf{x}^* - \mathbf{x}^0) = 0, \quad (\text{B.12})$$

which can be rearranged and solved for  $\mathbf{x}^*$ . In order to find a useful solution to the ill-posed problem, the analyst first generates a set of solutions with different values of  $\|\mathbf{x}\|_2$  and  $\|\boldsymbol{\delta}\|_2$  by varying the values of  $\beta_i$ ,  $i = 0 \dots \gamma$ , and then selects the one that offers the best compromise between solution accuracy and smoothness.

### B.3.3 Iterative Conjugate Gradient Regularization

The iterative conjugate gradient regularization technique is derived from an application of the conjugate gradient minimization technique used to solve linear systems of equations.

As shown in the previous section, it is possible to rewrite a system of linear equations as a minimization problem; in that case, it was implied that if  $\mathbf{A}$  were well-conditioned, the system of linear equations could be solved by minimizing the magnitude of the residual vector,  $F(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ ; since this objective function is a quadratic, there exists one global minimum,  $\mathbf{x}^*$ , that also solves the system of linear equations.

If  $\mathbf{A}$  is symmetric and positive-definite, a more convenient quadratic objective function is derived by setting the gradient equal to the residual vector,  $\nabla F(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$ , and noting that at  $\mathbf{x}^*$ ,  $\boldsymbol{\delta} = \nabla F(\mathbf{x}^*) = 0$ . Integrating with respect to  $\mathbf{x}$  yields a different objective function,

$$F(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x}. \quad (\text{B.13})$$

It can be shown that  $F(\mathbf{x})$  is a quadratic function having one global minimum,  $\mathbf{x}^*$ , that solves  $\mathbf{Ax} = \mathbf{b}$ . (Problems where  $\mathbf{A}$  is not symmetric are treated by solving  $\mathbf{A}^T\mathbf{Ax} = \mathbf{Ab}$ .) The conjugate-gradient minimization method, which is presented in detail in Appendix A, is particularly well suited to solve this problem; if  $\mathbf{A}$  is an  $n \times n$  matrix with  $n$  distinct eigenvalues, then the conjugate gradient method minimizes  $F(\mathbf{x})$  in  $n$  steps if exact arithmetic is used. This is accomplished by generating a set of mutually-conjugate “non-interfering” search directions, so that minimization in one direction is not “spoiled” by minimizing in subsequent directions.

As previously stated,  $\mathbf{A}$  must be positive-definite for the conjugate gradient method (or any gradient-based minimization method) to converge to  $\mathbf{x}^*$ , which occurs only when all the eigenvalues of  $\mathbf{A}$  are strictly positive. The necessity of this criterion can be seen by considering the geometric interpretation of the  $i^{th}$  eigenvalue,  $\mu_i$ , which is the amount by which the objective function increases when the solution is varied from  $\mathbf{x}^*$  by one unit in the direction of the  $i^{th}$  eigenvector,  $\lambda_i$ . The eigenvalues can also be used to define the condition number of  $\mathbf{A}$ ,

$$Cond(\mathbf{A}) = \frac{\mu_{\max}}{\mu_{\min}}, \quad (\text{B.14})$$

which is analogous to the condition number defined using the singular values in Eq. (B.5).

If all the eigenvalues are positive and of approximately the same order of magnitude, then  $\mathbf{A}$  is well-conditioned and there exists a unique solution to  $\mathbf{Ax} = \mathbf{b}$ ,  $\mathbf{x}^*$ . In such cases,  $F(\mathbf{x})$  is said to have a strong global minimum. For example, consider the system of linear equations given by

$$\mathbf{A} = \begin{bmatrix} 5 & 3 \\ 3 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -4 \\ -3 \end{bmatrix}. \quad (\text{B.15})$$

The corresponding objective function is found by substituting  $\mathbf{A}$  and  $\mathbf{b}$  into Eq. (B.13), and is shown in Fig. B.4 (a). In this example,  $\mathbf{A}$  has eigenvalues  $\mu_1 = 7.54$  and  $\mu_2 = 1.46$ , and the condition number equals 5.17. Consequently, there exists a strong global minimum  $\mathbf{x}^* = \{-0.67, -0.24\}^T$ .

In cases where  $\mu_{\min} \ll \mu_{\max}$ ,  $\mathbf{A}$  is ill-conditioned (as indicated by the large condition number), and there will be a set of solutions in the vicinity of  $\mathbf{x}^*$  that almost satisfy  $\mathbf{Ax} = \mathbf{b}$ . This situation is demonstrated by plotting the objective function generated from

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 2 & 1.05 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -4 \\ -2 \end{bmatrix}, \quad (\text{B.16})$$

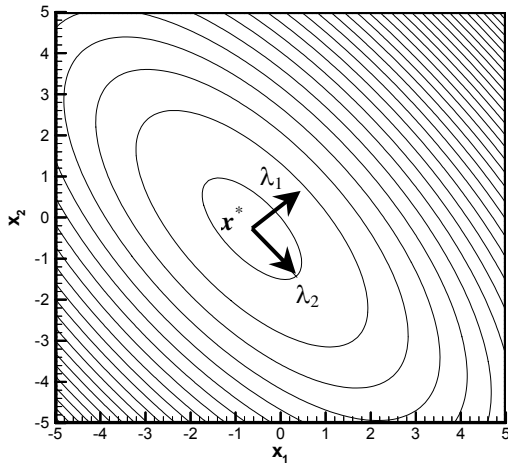
which is shown in Fig. B.4 (b). As before, there a strong global minimum  $\mathbf{x}^* = \{-1, 0\}^T$ , but unlike the previous example the first eigenvalue,  $\mu_1 = 5.01$ , is much larger than the second eigenvalue,  $\mu_2 = 0.04$ , and  $\mathbf{A}$  has a condition number of 125.25. The long, narrow valley leading up to the global minimum makes this objective function topography difficult to minimize, and any solution along the line defined by  $\{\mathbf{x} \in \mathbf{x}^* + \alpha \lambda_2\}$  in the vicinity of  $\mathbf{x}^*$  solves Eq. (B.16) with a small residual. (This is why there is a direct correlation between the condition number of  $\mathbf{A}$  and the number of iterations required to solve  $\mathbf{Ax} = \mathbf{b}$  using iterative linear solvers like Gauss-Seidel, which are derived from gradient-based minimization techniques.)

Finally, if  $\mathbf{A}$  is rank deficient with a nullity of  $k$ , then  $k$  of the eigenvalues equal zero and an infinite set of solutions minimize  $F(\mathbf{x})$ . In this case,  $\mathbf{A}$  is said to be semi-

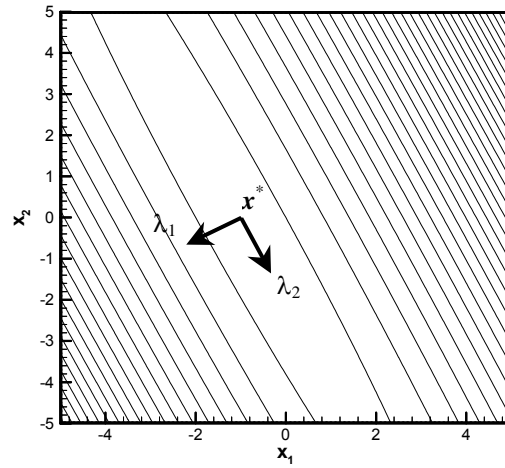
definite (and singular). Such an example is shown in Fig. B.4 (c), which is generated from

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -4 \\ -2 \end{bmatrix}. \quad (\text{B.17})$$

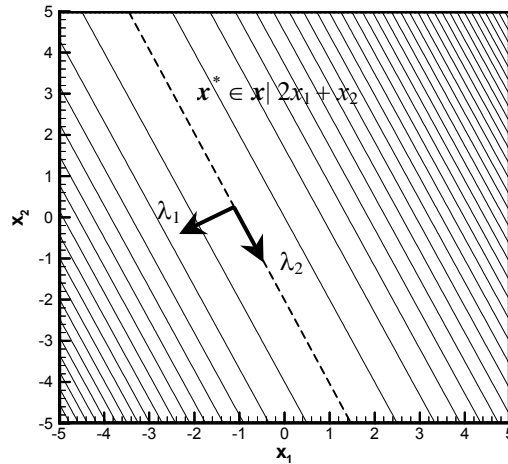
In this example,  $\mu_1 = 5$  and  $\mu_2 = 0$ , so the condition number of  $\mathbf{A}$  is infinite. Unlike the previous two examples, the objective function has a weak global minimum, and an infinite set of solutions defined by  $\{\mathbf{x}^* = \mathbf{x} \mid 2x_1 + x_2 = 0\}$  minimizes  $F(\mathbf{x})$ .



(a)  $\mu_1 = 7.54, \mu_2 = 1.46, \text{Cond}(\mathbf{A}) = 5.17$



(b)  $\mu_1 = 5.01, \mu_2 = 0.04, \text{Cond}(\mathbf{A}) = 125.25$



(c)  $\mu_1 = 5, \mu_2 = 0, \text{Cond}(\mathbf{A}) = \infty$

Figure B.4: Objective function topographies corresponding to cases where  $\mathbf{A}$  is (a) well posed, (b) ill-posed, and (c) singular.



In cases where  $\mathbf{A}$  is either ill-posed or degenerate, conjugate gradient minimization is used as a regularization method rather than a linear solver. This is possible because the conjugate gradient method minimizes  $F(\mathbf{x})$  in the directions corresponding to the largest eigenvalues first, and the minimizations involving the smallest eigenvalues (that cause the algorithm to fail) occur in the last few iterations. If minimization starts from  $x_i^0 = 0, i = 1 \dots n$ , which is presumably located far away from  $\mathbf{x}^*$ ,  $\|\mathbf{x}^0\|_2 = 0$  and  $\|\boldsymbol{\delta}^0\|_2 = \|\nabla F(\mathbf{x}^0)\|_2$  will be very large. In subsequent iterations,  $\mathbf{x}^k$  gets closer to  $\mathbf{x}^*$  and  $\|\mathbf{x}^k\|_2$  increases in a monotonic way while  $\|\nabla F(\mathbf{x}^k)\|_2$  and  $\|\boldsymbol{\delta}^k\|_2$  decrease in a monotonic way. In the vicinity of  $\mathbf{x}^*$ , the eigenvalues corresponding with the search directions are very small or zero; in the former case  $\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2$  becomes very large, while in the latter the minimization algorithm fails completely. In this way, the iteration number acts as a heuristic parameter to control the degree of regularization.

The conjugate gradient iterates can also be interpreted as the solutions to a set of well-posed problems that are in some way related to the original ill-posed problem. If  $k$  is small, an objective function with a global minimum at  $\mathbf{x}^k$  is very different from the one generated using the ill-conditioned set of equations; consequently, while  $\|\mathbf{x}^k\|_2$  may be small,  $\|\boldsymbol{\delta}^k\|_2$  is quite large. On the other hand, for large values of  $k$ , the objective function with a global minimum at  $\mathbf{x}^k$  closely matches the original objective function topography, so  $\|\boldsymbol{\delta}^k\|_2$  is quite small while  $\|\mathbf{x}^k\|_2$  is large.

## B.4 Demonstration of Regularization Methods

The three regularization methods discussed in the previous section are demonstrated by using each of them to solve the inverse radiant enclosure design problem shown in Fig. B.5. The radiant enclosure consists of a design surface, two adiabatic surfaces, and a heater surface. The design surface, located on the bottom, has an emissivity of  $\varepsilon_{DS} = 0.5$  and the heater surface at the top of the enclosure has an emissivity  $\varepsilon_{HS} = 0.9$ .

The objective of the design problem is to find a continuous heat flux distribution over the heater surface that produces a uniform emissive power  $E_{DS} = 0 \text{ W/m}^2$  and a uniform heat flux  $q_{sDS} = -1 \text{ W/m}^2$  over the design surface. This is done using the TSVD, Tikhonov, and iterative conjugate gradient regularization methods to solve an ill-conditioned system of linear equations generated using the procedure outlined in Chapter 2.

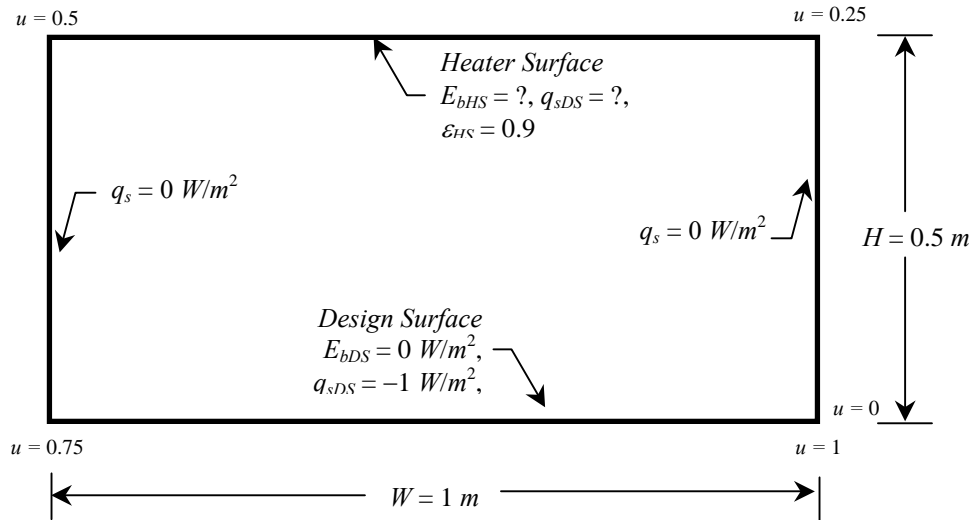


Figure B.5: Example inverse radiant enclosure design problem.

A total of 640 elements was chosen to solve the problem based on the results of the grid refinement study shown in Fig. B.6, which was carried out by solving the forward problem with a uniform heat flux distribution of  $q_{sHS} = 1 \text{ W/m}^2$  over the heater surface and  $E_{bDS} = 0 \text{ W/m}^2$  over the design surface. Of these, 160 elements were located on the design surface, 160 were located on the heater surface, and the remaining 320 were on the adiabatic surface. Unlike the example in Chapter 2 where large heaters having uniform heat flux distributions were located on the heater surface, a continuous heat flux distribution over the heater surface is modeled by approximating the heat flux over each finite area element as constant. This resulted in a system of 480 radiosity equations corresponding to elements on the adiabatic and heater surfaces, which contained a total of 640 unknown radiosity and heat flux values. An energy balance equation similar to Eq. (2.36) is added to reduce the nullity of  $\mathbf{A}$  from 160 to 159.

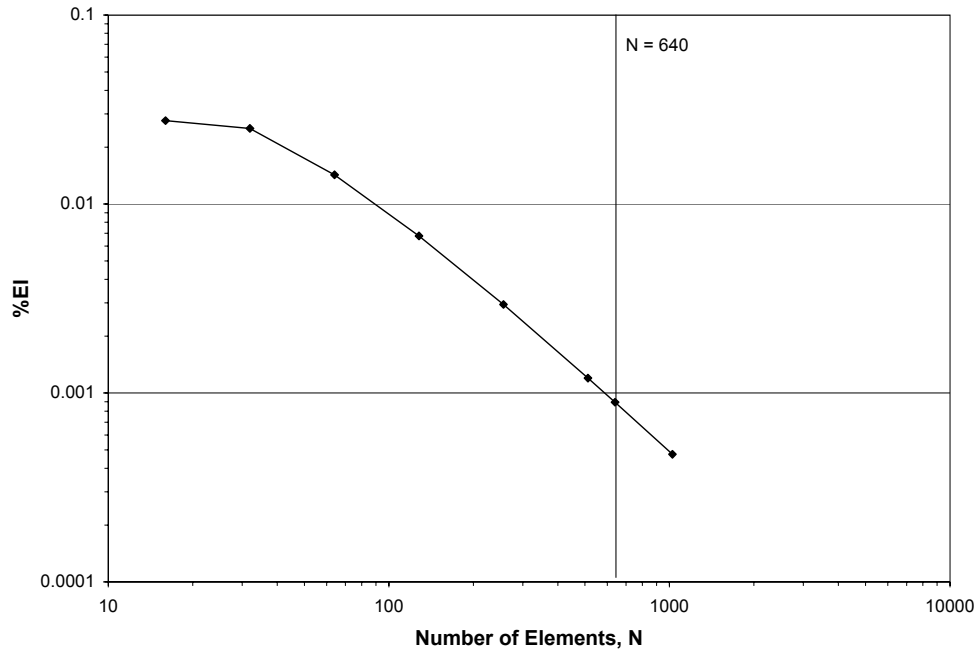


Figure B.6: Grid refinement study.

#### B.4.1 Truncated Singular Value Decomposition Regularization

The first step of this method is to perform the singular value decomposition of matrix **A**; the singular values are reordered and plotted in Fig. B.7. Exactly 159 of the singular values equal zero due to the nullity of **A**.

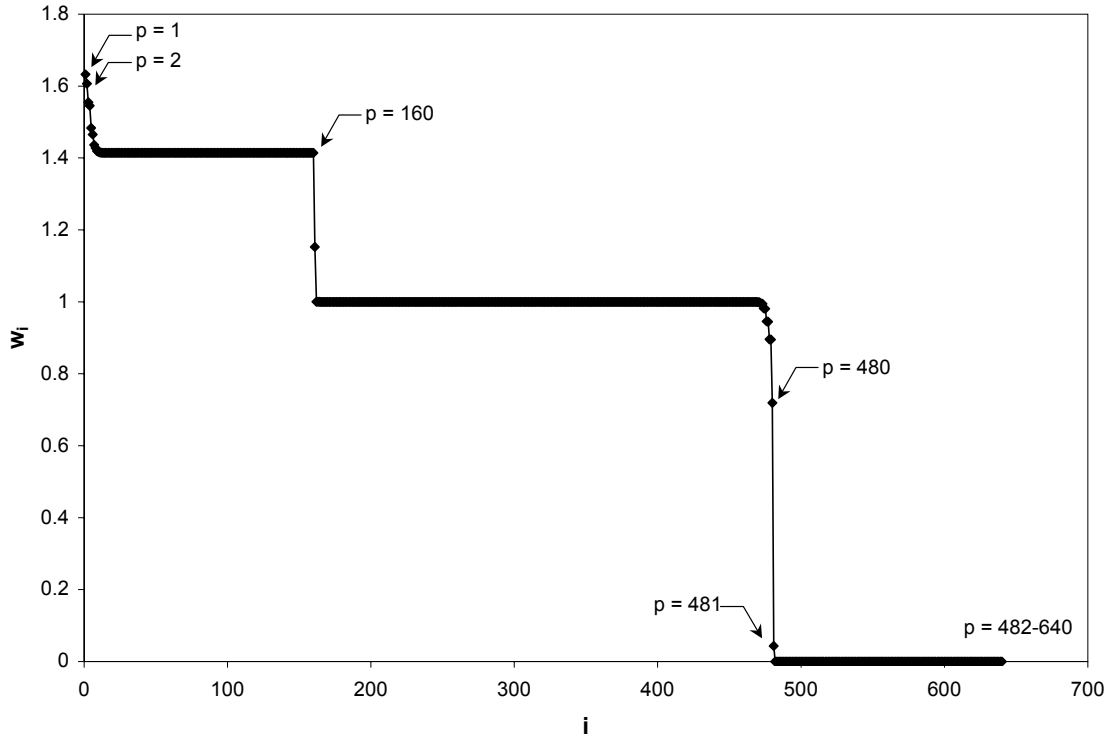


Figure B.7: Singular values of the **A** matrix.

Once this has been done, the analyst must then determine the degree of regularization that should be applied to the problem by choosing the number of singular values to truncate in Eq. (B.6). In order to do this, it is useful to first plot the “L-curve” shown in Fig. B.8, which indicates the degree of solution accuracy and smoothness, measured by  $\|\delta\|_2$  and  $\|x\|_2$  respectively, that result from varying the number of singular

values,  $p$ , used to compute the solution. (In this type of regularization, the L-curve is inexpensive to evaluate because the same singular value decomposition is used to generate each solution.)

Points on left-hand side of this curve correspond with very accurate but noisy solutions that are found by retaining most of the singular values. On the right-hand side of the curve, smooth but inaccurate solutions are found by truncating most of the singular values. In many (but by no means all) cases, the best compromise between accuracy and smoothness is found in the middle of the curve, near the apex of the “L.”

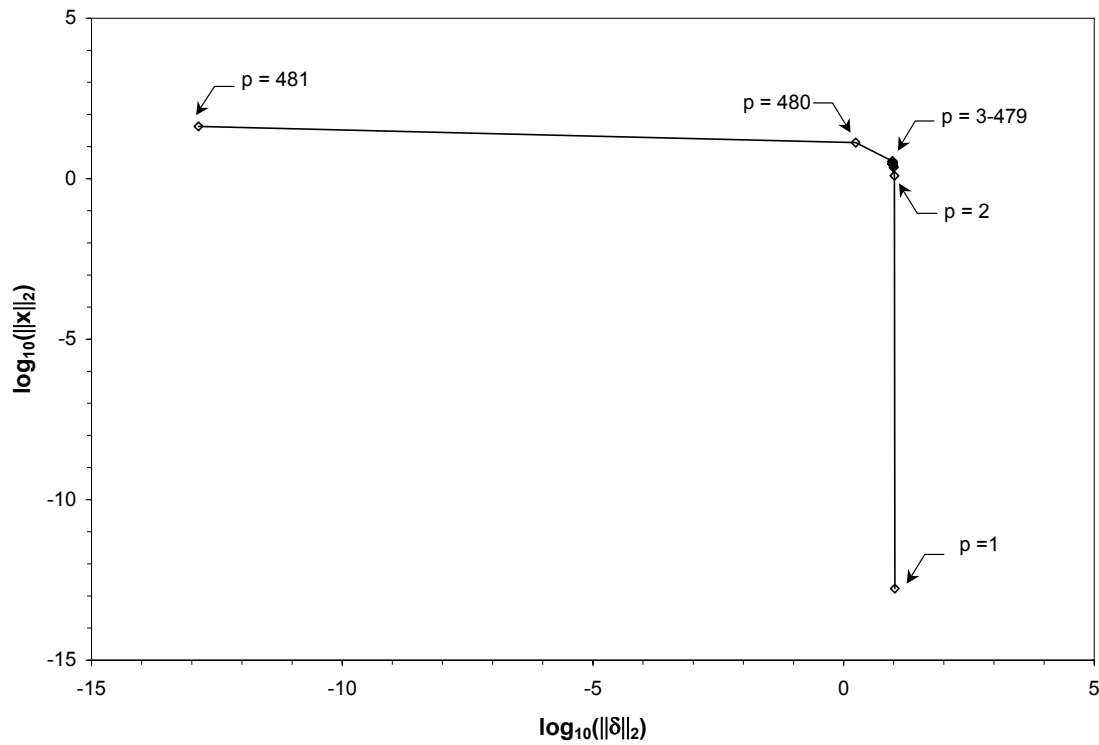


Figure B.8: L-Curve formed using TSVD regularization.

Solutions obtained using  $p = 1, 2, 160, 480$ , and 481 singular values are shown in Figs. B.9 and B.10. Figure B.9 shows the heat flux and emissive power distribution calculated over the heater surface, and Fig. B.10 shows the corresponding heat flux distributions over the design surface found by substituting the inverse solutions into the forward problem and specifying the emissive power distribution over the design surface. The corresponding solution and residual norms are shown in Table B.1.

From examining Fig. B.9 and the solution and residual norms in Table B.1, it is clear that the magnitude of the heat flux distribution over the heater surface becomes smaller as more singular values are truncated, while the norm of the residuals become larger. In this case, however, the only useful solution is found by truncating just the null singular values ( $p = 481$ ), since further truncation results in a negative heat flux distribution over the heater surface. This is also clear from Fig. B.10, which shows that the only heat flux distribution over the design surface that closely matches the desired distribution is the obtained using the heater inputs calculated using  $p = 481$  singular values.

Number of singular values used, $p$	$\ \delta\ _2$	$\ \mathbf{x}\ _2$
1	10.543	$1.691 \times 10^{-13}$
2	10.358	1.226
160	9.580	2.877
480	1.738	13.303
481	$1.358 \times 10^{-13}$	42.240

Table B.1: Residual and solution norms of different TSVD solutions.

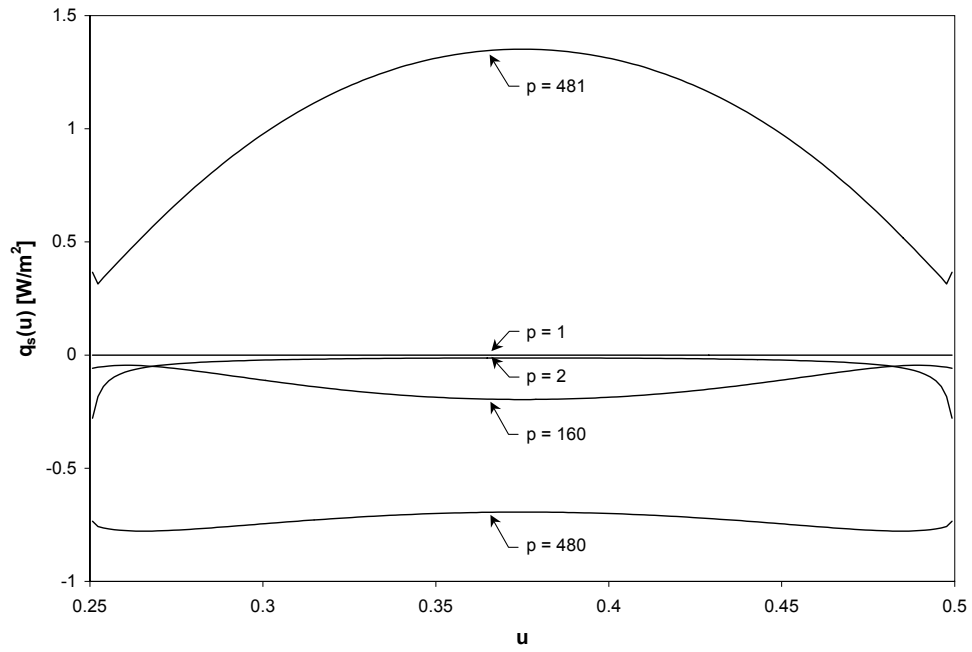


Figure B.9: Heat flux distributions over the heater surface obtained through TSVD.

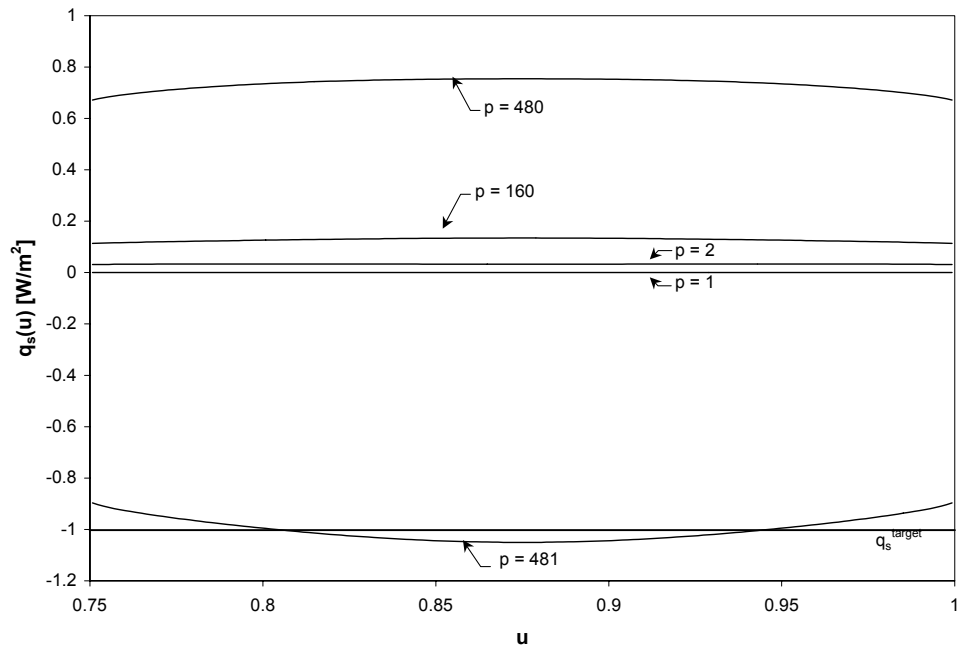


Figure B.10: Heat flux distributions over the design surface obtained through TSVD.

#### B.4.2 Tikhonov Regularization

Tikhonov regularization was next applied to solve the ill-conditioned governing equations. A second-order scheme was chosen based on the observation that radiosity and heat flux distributions over each surface usually have a parabolic shape for this particular radiant enclosure geometry type. In this application,  $\rho_0 = -0.001$ ,  $\rho_1 = 0$ ,  $\rho_2 = 1$ , and different levels of regularization were achieved by changing the variable  $c$  in Eq. (B.11).

As in the previous method, when performing Tikhonov regularization it is useful to first generate the L-curve in order to assess what amount of regularization is likely to produce a solution that is an acceptable compromise between accuracy and smoothness. In this method, however, every different value of  $c$  generates a unique set of linear equations that must be solved individually (through LU decomposition, in this application), and consequently the Tikhonov L-curve requires substantially more computational effort to plot than the TSVD L-curve.

Figure B.11 shows the L-curve formed by varying the regularization parameter over the range  $1 \times 10^{-7} \leq c \leq 10^7$ . Points on the left-hand side of the curve correspond to very small values of  $c$ ; since the systems of linear equations formed by rearranging Eq. (B.12) is very similar to the original ill-conditioned system of equations, the residual norms are quite small but the solution norms are very large. At a critically small value of  $c$ , (in this case  $c^{crit} \approx 1 \times 10^{-8}$ ), the system of linear equations are so close to being singular that truncation and round-off errors in the floating-point arithmetic cause a division-by-zero error during the decomposition process.



Points on the right-hand side of the curve are generated using very large values of  $c$ ; in these cases, far more emphasis is placed on minimizing the objective function defined in Eq. (B.9) rather than minimizing the residual vector of the original set of ill-conditioned equations. Accordingly, in these cases the solution norms are very small but the residual norms are quite large. As  $c$  approaches infinity,  $\|\mathbf{x}\|$  goes to zero and  $\|\boldsymbol{\delta}\|_2$  approaches  $\|\mathbf{b}\|_2$  asymptotically.

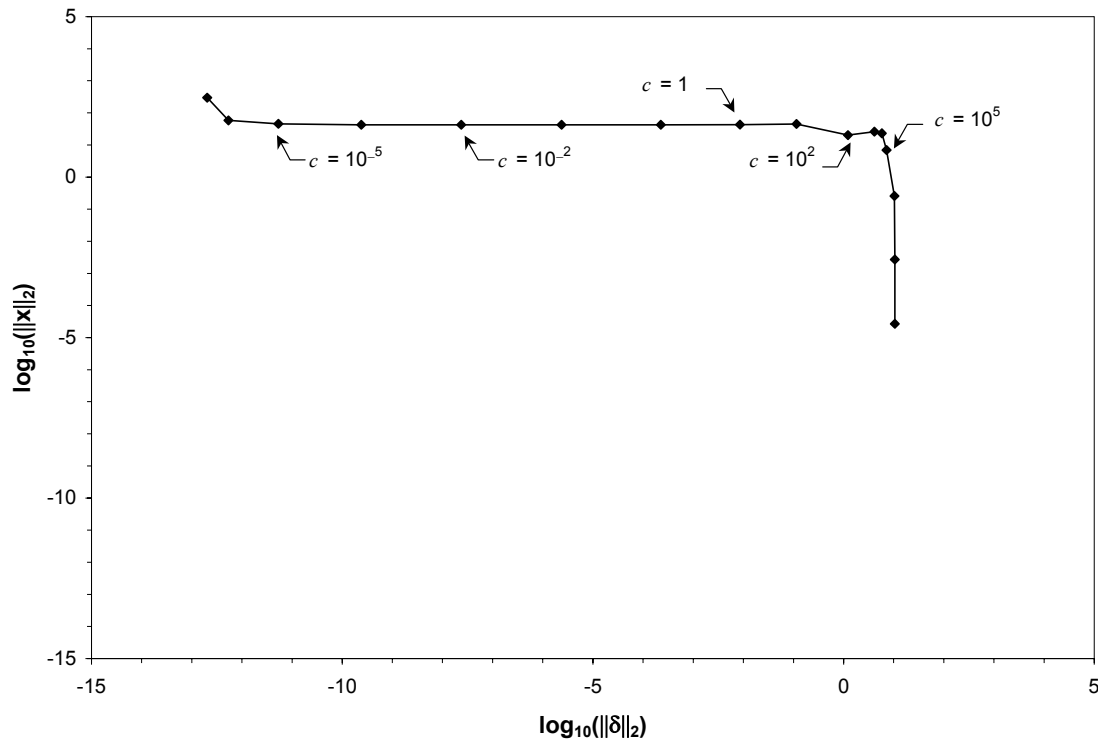


Figure B.11: L-curve formed using Tikhonov regularization.

Solutions obtained using  $c = 10^{-5}$ ,  $10^{-2}$ ,  $1$ ,  $10^2$ , and  $10^5$  are shown in Figs. B.12 and B.13, while the norms of the corresponding solution and residual vectors are included in Table B.2. Increasing  $c$  both decreases the magnitude of the solution and normalizes the heat flux distribution. As shown in Fig. B.12 and Table B.2, at  $c = 10^{-5}$  the solution norm is quite large because the corresponding set of linear equations is ill-conditioned.

(The asymmetry in the heat flux distribution is caused by round-off and truncation errors that occur during LU decomposition.) The magnitude of the heat flux distribution is significantly lower for  $c = 10^{-2}$  and  $c = 1$ , and the distribution becomes smoother.

Increasing  $c$  also increases the solution residual, however. As shown in Fig. B.13, while the solutions obtained using  $c = 10^{-2}$  and  $c = 1$  are both very close to the desired heat flux distribution over the design surface and have correspondingly small residuals, further increasing  $c$  drives the heat flux distribution away from the desired distribution as shown in the cases of  $c = 10^2$  and  $c = 10^5$ . The magnitudes of the residual vector are quite large for these last two cases, which indicates that the original ill-conditioned set of linear equations is not enforced. In fact, the heat flux values over the heater surface are less than the heat flux distribution over the design surface, so these solutions are clearly non-physical.

Regularization parameter, $c$	$\ \delta\ _2$	$\ \mathbf{x}\ _2$
$10^5$	7.204	6.970
$10^2$	1.221	20.207
1	$8.551 \times 10^{-3}$	43.137
$10^{-2}$	$2.375 \times 10^{-6}$	42.677
$10^{-5}$	$5.321 \times 10^{-12}$	45.776

Table B.2: Residual and solution norms of different Tikhonov solutions.

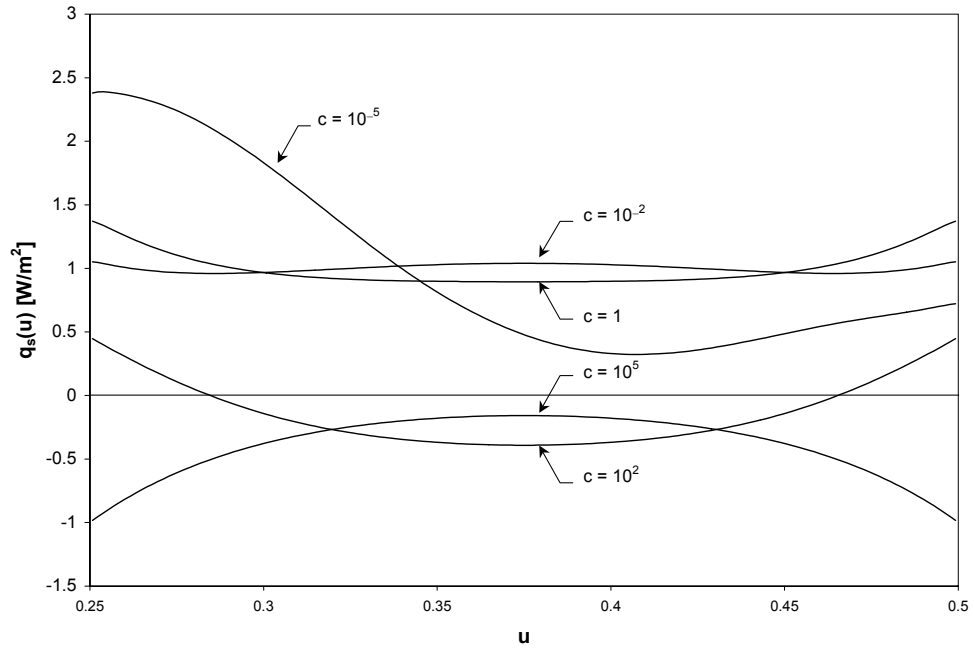


Figure B.12: Heat flux distributions over the heater surface obtained through Tikhonov regularization.

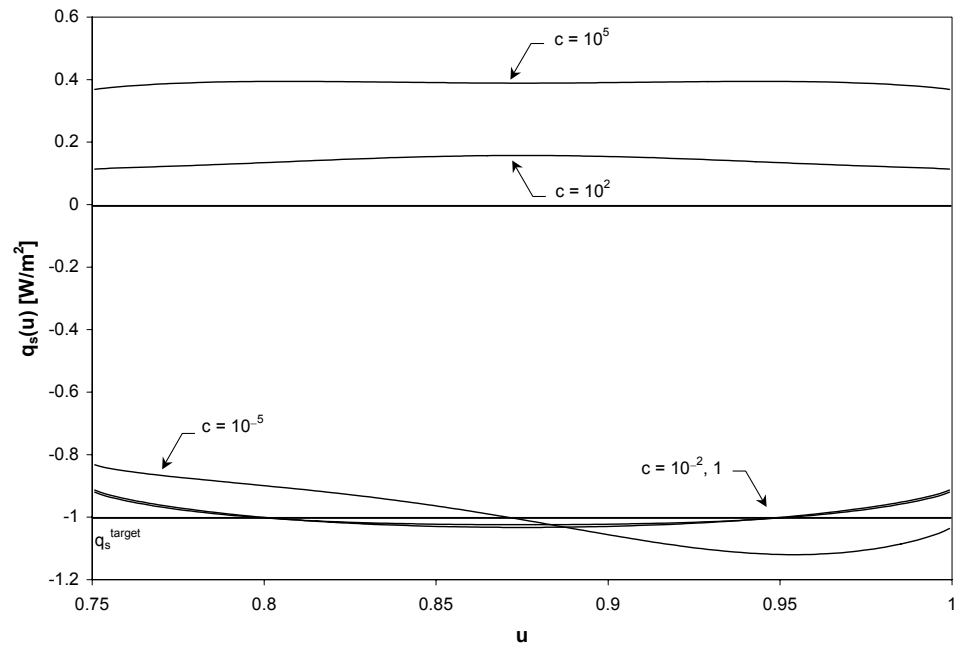


Figure B.13: Heat flux distributions over the design surface obtained through Tikhonov regularization.

### B.4.3 Iterative Conjugate Gradient Regularization

Finally, the ill-conditioned system of linear equations was solved using iterative conjugate gradient regularization. In this application, minimization was started at  $x_i^0 = 2, i = 1 \dots n$ .

As before, the first step in the solution process is to plot the magnitudes of the residual and solution vectors obtained at different levels of regularization. Instead of an L-curve, however, in this regularization method  $\|\delta^k\|_2$  and  $\|\mathbf{x}^k\|_2$  are both plotted as functions of the iteration number, as shown in Fig. B.14. After the first conjugate gradient iteration the solution norm grows slightly and the residual becomes smaller as  $\mathbf{x}^k \rightarrow \mathbf{x}^*$ , for  $1 \leq k \leq 7$ ; at these values, the Hessian of the objective function defined in Eq. (B.13) is well-conditioned, so both  $\mathbf{p}^k$  and  $\alpha^k$  are bounded. At  $k = 8$ , however,  $\nabla F(\mathbf{x}^k)$  becomes ill-conditioned as some of the eigenvalues approach zero. This causes the conjugate gradient iterates to fail, and both  $\|\mathbf{x}^k\|_2$  and  $\|\delta^k\|_2$  become very large. (This situation is akin to reaching the bottom of a very shallow valley or trough when minimizing an objective function having two variables). Thus, unlike the previous two methods where  $\|\delta\|_2$  varied in a monotonic way with respect to the regularization heuristics (the number of truncated singular values in TSVD and  $c$  in Tikhonov regularization), in conjugate gradient regularization  $\|\delta\|_2$  first decreases and then increases with increasing iteration number.

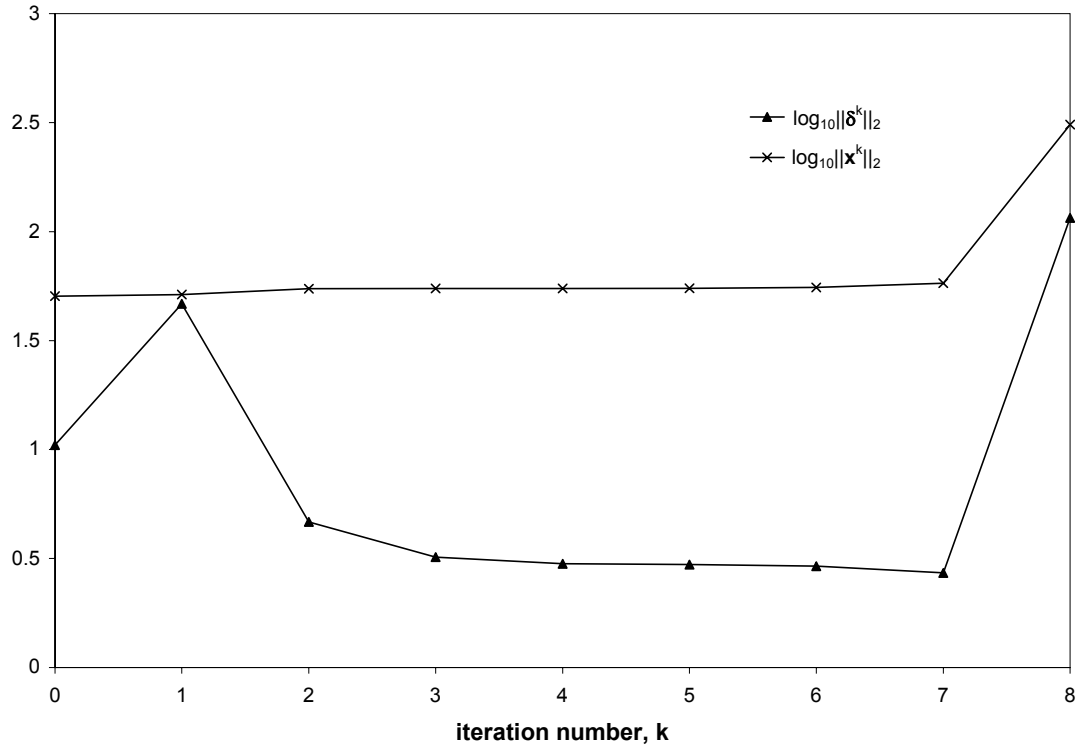


Figure B.14: Plot of solution and residual norms as a function of iteration number for iterative conjugate gradient regularization.

The heat flux distributions over the heater and design surfaces are shown in Figs. B.15 and B.16, while the magnitudes of the solution and residual vectors are included in Table B.3. As the iteration number increases from  $k=1$  to  $k=3$ , the magnitude of the heat flux distribution over the heater surface becomes smaller, as shown in Fig. B.15, and the heat flux distribution on the design surface steadily approaches the desired heat flux distribution, as shown in Fig. B.16. Based on the residual vector magnitudes shown in Table B.3, it is clear that for these three iterations  $x^k$  is steadily approaching  $x^*$  and the conjugate gradient iterates are working. The solutions obtained using  $k=3, 4$ , and  $5$  are

very similar, and  $\|\delta^k\|_2$  and  $\|\mathbf{x}^k\|_2$  change very little with increasing iterations. For  $k = 5$ ,  $k = 6$ , and  $k = 7$ , the magnitude of the residual vector continues to decrease and the solution norm grows with continuing  $k$ . The heat flux distributions over the heater surface increase in magnitude and the heat flux distribution over the design surface moves away from the desired distribution for these cases, as shown in Figs. B.15 and B.16. While the magnitude of the residual vector is decreasing for these three cases, the  $L_2$  norm of the subspace of  $\delta$  corresponding to the radiosity equations (i.e. excluding the energy balance equation) is actually increasing. For  $k > 7$ ,  $\nabla^2 \mathbf{F}(\mathbf{x}^k)$  becomes ill conditioned, as indicated by the large values of  $\|\delta^k\|_2$  and  $\|\mathbf{x}^k\|_2$  shown in Table B.3.

Iteration number, $k$	$\ \delta^k\ _2$	$\ \mathbf{x}^k\ _2$
0	10.496	50.596
1	46.610	51.375
2	4.659	54.671
3	3.213	54.796
4	2.992	54.857
5	2.970	54.898
6	2.917	55.456
7	2.719	58.002
8	115.667	309.411
9	3871.09	1486.21

Table B.3: Residual and solution norms of different conjugate gradient iterates

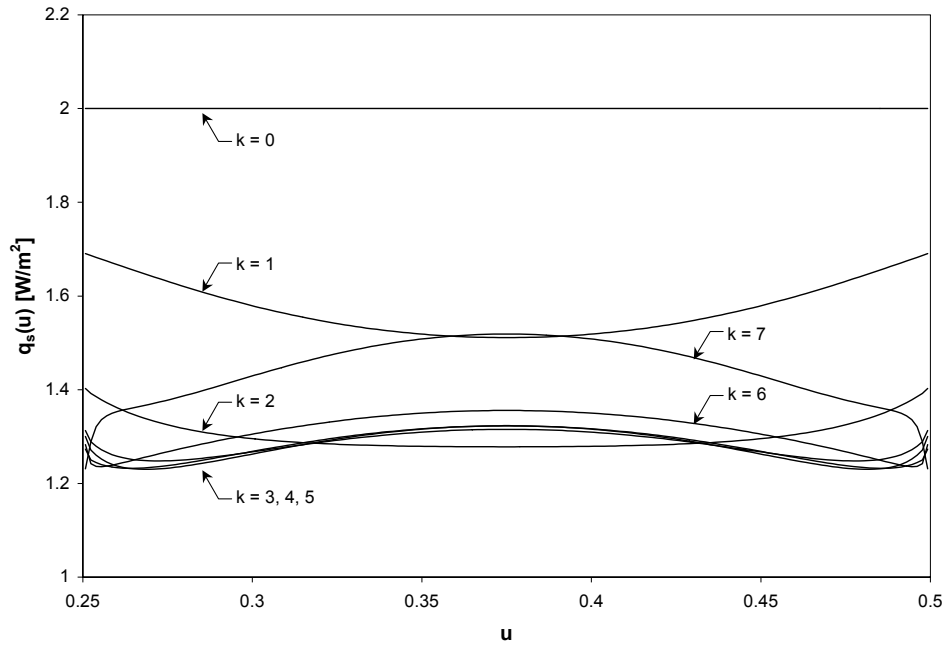


Figure B.15: Heat flux distributions over the heater surface obtained through iterative conjugate gradient regularization.

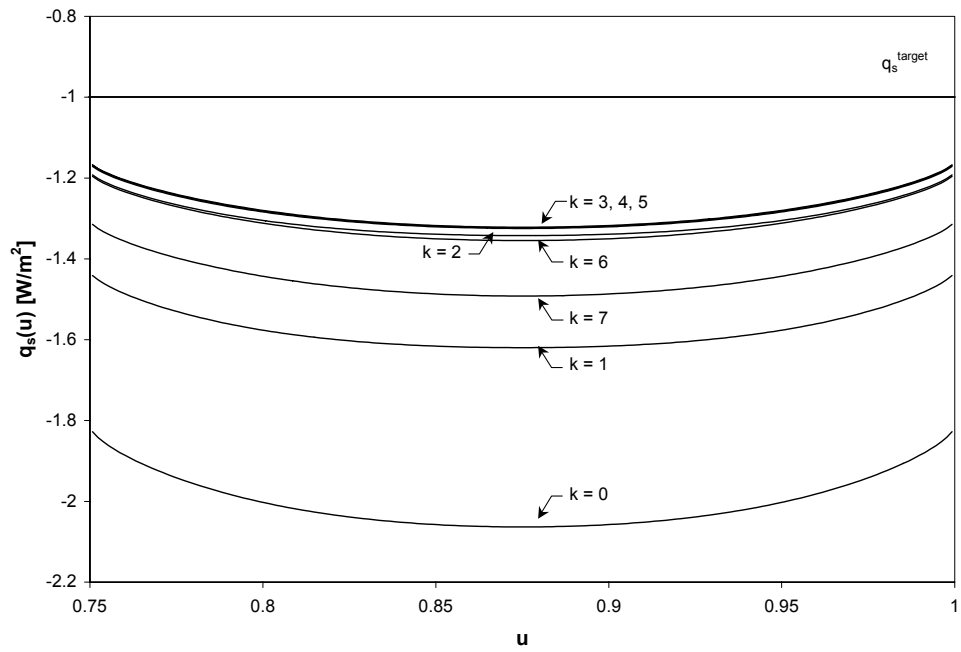


Figure B.16: Heat flux distributions over the heater surface obtained through iterative conjugate gradient regularization.

#### 3.4.4 Comparison of Regularization Methods

The three regularization methods are evaluated and compared based on three criteria: the amount of information that each method provides about the ill-posed problem, the solution quality, and the relative ease of implementation.

Of the three methods examined here, the TSVD method provides the most insight into the nature of the ill-posed problem. By performing the singular value decomposition and plotting the singular value spectra shown in Fig. B.7, the nullity of  $\mathbf{A}$  is easily found by counting the number of singular values equal to zero. More importantly, the existence of different solution modes that satisfy the original ill-posed problem with a small residual can be anticipated if there are many small but non-zero singular values. Such *a priori* information is not provided by the Tikhonov and iterative conjugate gradient methods.

The best solutions obtained using each regularization method, i.e. those that produce heat flux distributions over the design surface that most closely match the desired distribution, are plotted in Fig. B.17. The TSVD solution was found using  $p = 481$  singular values, the Tikhonov solution used  $c = 10^{-2}$ , and the iterative conjugate gradient solution was found with  $k = 5$  iterations. The heat flux distributions over the design surface found by substituting the inverse solutions into the forward problem are plotted in Fig. B.18. The solution obtained using Tikhonov regularization most accurately satisfies the desired heat flux distribution over the design surface, followed by the TSVD solution and the iterative conjugate gradient solution. From this perspective, one of the main advantages of Tikhonov regularization is that the heuristic parameter is



continuous and can be varied to produce a large set of well-posed problems related to the original ill-posed problems, greatly increasing the likelihood of finding a solution that is both sufficiently accurate and regular. The regularization parameters of the other two methods, on the other hand, are discrete and produce a finite set of potential solutions; the size of the TSVD solution set is limited by the number of unique singular values, while the iterative conjugate gradient solution set is limited by the number of convergent conjugate gradient iterations. Accordingly, there is somewhat less likelihood of finding a useful solution from these smaller solution sets.

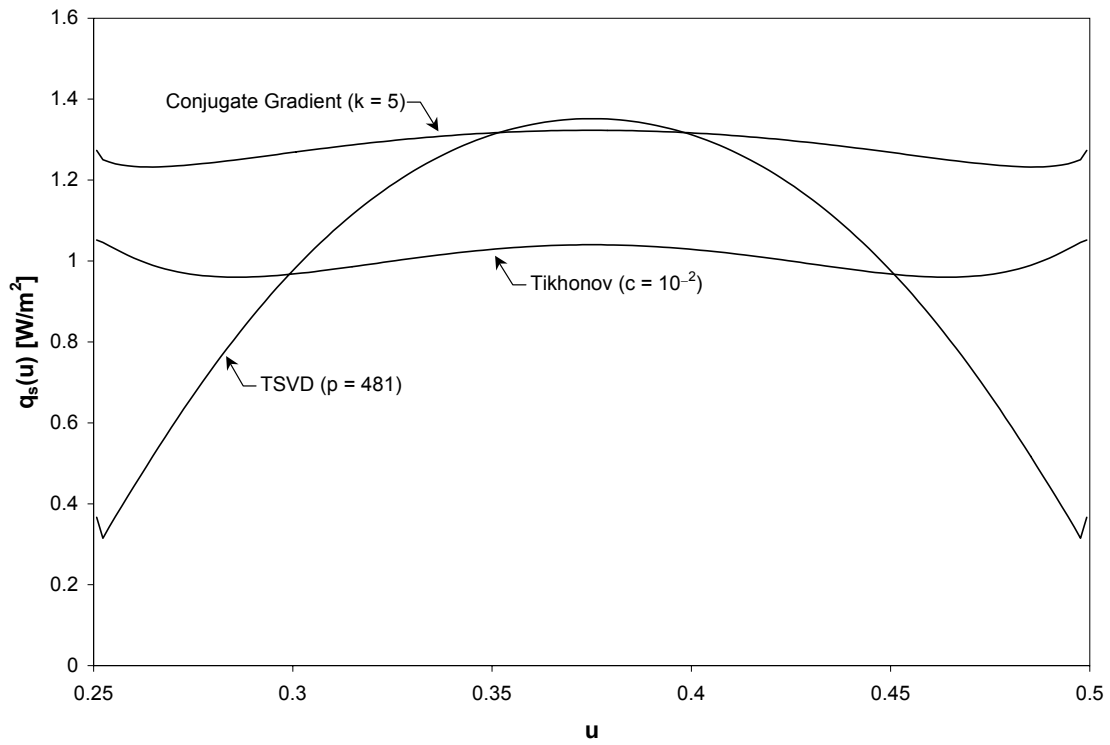


Figure B.17: Heat flux distributions over the heater surface that most closely satisfy the desired distribution over the design surface.

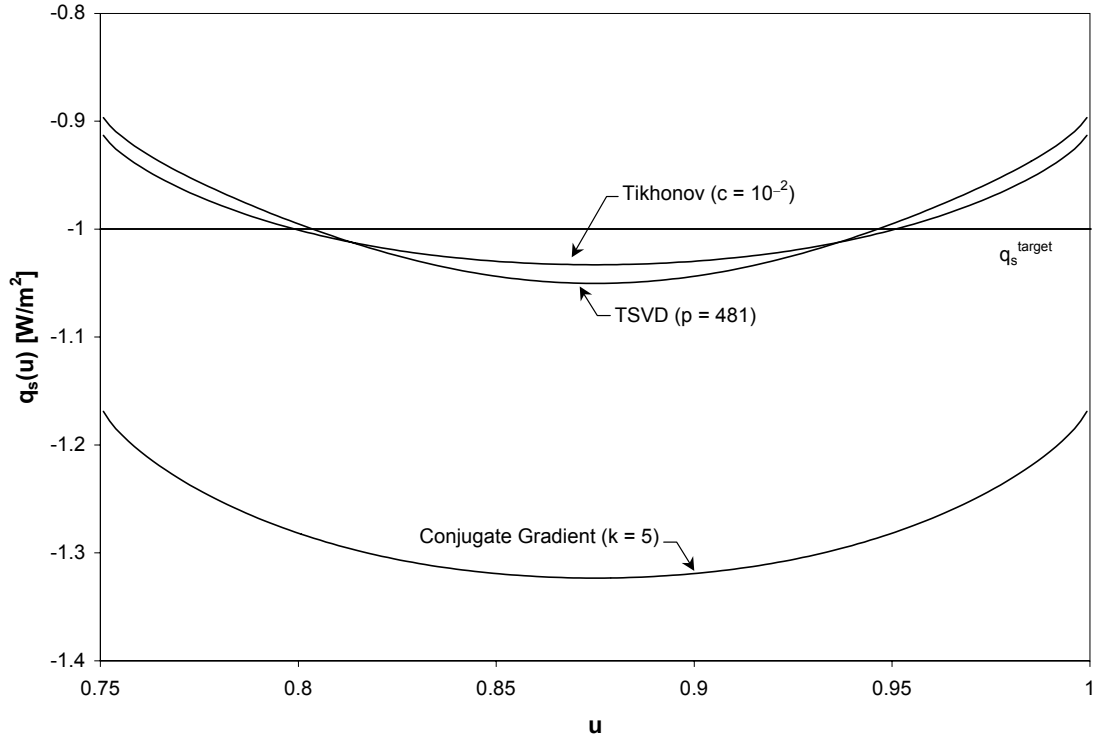


Figure B.18: Heat flux distributions over the design surface corresponding to the solutions shown in Fig. B.17.

The iterative conjugate gradient method is by far the most computationally efficient method, since each regularization iteration involves only matrix and vector multiplication. The TSVD method, on the other hand, relies on the singular value decomposition of  $\mathbf{A}$ , demanding approximately the same amount of CPU time and memory storage as each LU decomposition used in Tikhonov regularization. Whereas the TSVD method requires only one singular value decomposition to calculate the entire set of regularized solution, however, when using Tikhonov regularization a new set of linear equations must be solved each time the regularization parameter is changed.

Moreover, because the Tikhonov regularization parameter is continuous and since it is difficult to predict what regularization parameter and smoothing matrix should be used, it is often necessary for the analyst to solve many different sub-problems until a useful solution to the ill-posed problem is identified. Consequently, Tikhonov regularization requires far most time and computational effort to carry out compared with the iterative conjugate gradient and TSVD methods.

## Bibliography

Alifanov, O. M., 1977, "Solution of an Inverse Problem of Heat-Conduction by Iterative Methods," *Journal of Engineering Physics*, **26**, 4, pp. 471-476.

Alifanov, O. M., 1995, *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems*, Begall House, New York, NY.

Ashdown, I., 1993, "Non-Imaging Optics Design Using Genetic Algorithms," *Journal of the Illuminating Engineering Society*, **23**, pp 12-21.

Balakrishnan, K. S., and Edgar, T. F., 2000, "Model-Based Control in Rapid Thermal Processing," *Thin Solid Films*, **365**, pp. 322-333.

Beck, J. V., Blackwell, B., St Clair, C. R., 1985, *Inverse Heat Conduction: Ill-posed Problems*, John Wiley and Sons, New York, NY.

Bertsekas, D. P., 1999a, *Nonlinear Programming*, 2<sup>nd</sup> Ed., Athena Scientific, Belmont MA, pp. 29-31.

Bertsekas, D. P., 1999b, *Nonlinear Programming*, 2<sup>nd</sup> Ed., Athena Scientific, Belmont MA, pp. 94-96.

Bertsekas, D. P., 1999c, *Nonlinear Programming*, 2<sup>nd</sup> Ed., Athena Scientific, Belmont MA, pp. 32.

Birkebak, R. C., Sparrow, E. M., Eckert, E. R. G., and Ramsey, J. W., 1964, "Effect of Surface Roughness on Total Hemispherical and Specular Reflectance on Metallic Surfaces," *Journal of Heat Transfer*, **C87**, pp. 429-435.

Björck, Å., 1996, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, PA, pp. 201-203.

Breedijk, T., Edgar, T. F., and Trachtenbergm I., 1994, "Model-Based Control of Rapid Thermal Processes," *Proc. American Control Conference*, Baltimore, MD.

Chen, J. Y., and Ryoji, H., 2000, "Combustion Process Optimization by Genetic Algorithms: Reduction of NO<sub>2</sub> Emission via Optimal Postflame Process," *Symposium (International) on Combustion*, **28**, 2, pp. 2483-2489.

Chen, Y. T., Chong, K. K., Bligh, T. P., Chen, L. C., Yunis, J., Kannan, K. S., Lim, B. H., Lim, C. S., Alias, M. A., Bidin, N., Aliman, O., Salehan, S., Shk. Abd. Rezan, S. A. H., Tam, C. M., and Tan, K. N., 2001, "Non-Imaging, Focusing Heliostat," *Solar Energy*, **71**, 3, pp. 155-164.

Cho, Y. M., and Gyugyi, P., 1997, "Control of Rapid Thermal Processing: A System Theoretic Approach," *IEEE Transactions on Control Systems Technology*, **5**, 6, pp. 644-653.

Choi, J. Y., and Do, H. M., 2001, "A Learning Approach of Wafer Temperature Control in a Rapid Thermal Processing System," *IEEE Transactions on Semiconductor Manufacturing*, **14**, 1, pp. 1-10.

Corana, A., Marchesi, M., Martini, C., and Ridella, S., 1987, "Minimizing Multimodal Functions of Continuous Variables with the Simulated Annealing Algorithm," *ACM Transactions on Mathematical Software*, **13**, 3, pp. 262-280.

Daun, K. J., and Hollands, K. G. T., 2001, "Infinitesimal Radiative Analysis Using Parametric Surface Representation, through NURBS," *Journal of Heat Transfer*, **123**, 2, pp. 249-256.

Daun, K. J., Howell, J. R., and Morton, D. P., 2003a, "Design of Radiant Enclosures using Inverse and Optimization Design Methodologies," *Inverse Problems in Engineering*, in press.

Daun, K. J., Howell, J. R., and Morton, D. P., 2003b, "Optimization of Industrial Processes involving Radiant Heating, through Nonlinear Programming," *Proc. IMECE 2003*, Washington D.C.

Daun, K. J., Howell, J. R., and Morton, D. P., 2003c, "Geometric Optimization of Radiant Enclosures through Nonlinear Programming," *Numerical Heat transfer Part B: Fundamentals*, **43**, 3, pp. 203-219.

Daun, K. J., Morton, D. P., and Howell, J. R., 2003d, "Geometric Optimization of Radiant Enclosures containing Specular Surfaces," *Journal of Heat Transfer*, **125**, 5, pp. 845-851.

Ebrahimi, S. A., Tortorelli, D. A., and Dantzig, J. A., 1997, "Sensitivity Analysis and Nonlinear Programming Applied to Investment Casting Design," *Applied Mathematical Modelling*, **21**, 2, pp. 113-123.

Ertürk, H., Ezekoye, O. A., and Howell, J. R., 2002, "The Application of an Inverse Formulation in the Design of Boundary Conditions for Transient Radiating Enclosures," *Journal of Heat Transfer*, **124**, 6, pp. 1088-1094.

Fabbri, G., 1998, "Heat Transfer Optimization in Finned Annular Ducts under Laminar Flow Conditions," *Heat Transfer Engineering*, **19**, 4, pp. 42-54.

Farin, G., 1997, *Curves and Surfaces for Computer Aided Design: A Practical Guide*, 4<sup>th</sup> Ed., Academic Press, San Diego, CA.

Fedorov, A. G., Lee, K. H., and Viskanta, R., 1998, "Inverse Optimal Design of the Radiant Heating in Materials Processing and Manufacturing," *Journal of Materials Engineering and Performance*, **7**, 6, pp. 719-726.

Forrest, S., 1993, "Genetic Algorithms: Principles of Natural Selection Applied to Computation," *Science*, 261, pp. 872-878.

França, F. H. R., and Goldstein, L., 1996, "Application of Zoning Method in Radiative Inverse Problems," *Proc. Brazilian Congress of Engineering and Thermal Sciences, ENCIT 96*, Florianopolis, Brazil.

França, F. H. R., Ezekoye, O. A., and Howell, J. R., 1999, "Two-Dimensional Inverse Heat Load Problem for Combined Radiation and Conduction," *Proc. 1999 IMECE*, Nashville, TN.

França, F. H. R., Ezekoye, O. A., and Howell, J. R., 2001, "Inverse Boundary Design Combining Radiation and Convection Heat Transfer," *Journal of Heat Transfer*, **123**, 5, pp. 884-891.

França, F. H. R., Howell, J. R., Ezekoye, O. A., and Morales, J. C., "Inverse Design of Thermal Systems," *in press for Advances in Heat Transfer*, J.P. Hartnett and T.F. Irvine, eds., **36**, Academic Press, 2002.

Gill, P. E., Murray, W., and Wright, M., 1986a, *Practical Optimization*, Academic Press, San Diego CA, pp. 119.

Gill, P. E., Murray, W., and Wright, M., 1986a, *Practical Optimization*, Academic Press, San Diego CA, pp. 107-111.

Gordon, J. M., 1993, "New High-Efficiency Uniform Illumination Devices based on Nonimaging Optics," *Proceedings of the SPIE, The International Society for Optical Engineering*, **1971**, pp. 368-377.

Güven, H. M., 1994, "Optimization of Parabolic Trough Collector Design for Varying Manufacturing Tolerances Using a Closed-Form Expression for Intercept Factor," *Journal of Solar Energy Engineering*, **116**, 3, pp. 164-166.

Gwak, K. W., and Masada, G., 2002, "Structural Analysis of Nonlinear Control Systems using Singular Value Decomposition," *Proc. 2002 IMECE*, New Orleans, LA.

Hadamard, J., 1923, *Lectures on Cauchy's Problem in Linear Differential Equations*, Yale University Press, New Haven, CT.

Hammersley, J. M., and Handscomb, D. C., 1992, *Monte Carlo Methods*, Chapman and Hall, London, pp. 14.

Hansen, P. C, 1998, *Rank Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM, Philadelphia, PA.

Han, S. Y, Maeng, J. S., and Dal-Hyun, Y., 2003, "Shape Optimization of Cutoff in a Multiblade Fan/Scroll; System using Response Surface Methodology," *Numerical Heat Transfer Part B: Fundamentals*, **43**, 1, pp. 87-98.

Harutunian, V., Morales, J. C., and Howell, J. R., 1995, "Radiation Exchange within an Enclosure of Diffuse-Gray Surfaces: The Inverse Problem," *Proc. ASME/AIChE 35<sup>th</sup> National Heat Transfer Conference*, Portland, OR.

Hiroyasu, H., Miao, H., Hiroyasu, T, Miki, M., Kamiura, J., and Watanabe, S., 2003, "Genetic Algorithms Optimization of Diesel Engine Emissions and Fuel Efficiency with Air Swirl, EGR, Injection Timing and Multiple Injections," *Proc. 2003 JSAE/SAE International Spring Fuels and Lubricants Meeting*, Yokohama.

Hollands, K. G. T., 1999, Class notes for ME 653, Thermal Radiation Heat Transfer, University of Waterloo.

Hosseini Savari, S. M., Howell, J. R., and Mansouri, S. H., 2003a, "Inverse Boundary Design Conduction-Radiation Problem in Irregular Two-Dimensional Domains," *Numerical Heat Transfer Part B: Fundamentals*, **44**, 3, pp.209-224.

Hosseini Savari, S. M., Howell, J. R., and Mansouri, S. H., 2003b, "A General Method for Estimation of Boundary Conditions over the Surface of Shields Surrounded by Radiating Enclosures," *Numerical Heat Transfer Part B: Fundamentals*, **44**, 1, pp. 25-43.

Huang, C. H., and Chen, H. M., 1999, "Inverse Geometry Problem of Identifying Growth of Boundary Shapes in a Multiple Region Domain," *Numerical Heat Transfer Part A: Applications*, **35**, 4, pp. 435-450.

Jones, M. R., 1999, "Inverse Analysis of Radiative Heat Transfer Systems," *Journal of Heat Transfer*, **121**, 2, pp. 481-484.

Kiefer J., and Wolfowitz, J., 1952, "Stochastic Estimation of the Maximum of a Regression Function," *Annals of Mathematical Statistics*, **23**, pp. 462-466.

Krishnaprakas, C. K., and Badari Narayana, K., 2003, "Heat Transfer Analysis of Mutually Irradiating Fins," *International Journal of Heat and Mass Transfer*, **46**, 5, pp 761-769.

Kudo, K, Kuroda, A., Eid, A., Saito, T., and Oguma, M., 1996, "Solution of the Inverse Radiative Load Problems by the Singular Value Decomposition," in *Radiative Transfer I: Proc. of the First International Symposium on Radiation Transfer*, M. Pinar Menguç ed., Bengell House, New York, NY, pp. 568-578.

Kushner, H. J., and Clark, D. S., 1978, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer, New York, NY.

Li, Q., Stephen, G. P., Querin, O. M., and Xie, Y. M., "Shape and Topology Design for Heat Conduction by Evolutionary Structural Optimization," *International Journal of Heat and Mass Transfer*, **42**, 17, pp. 3361-3371.

Liu, L. H., 2000, "Simultaneous Identification of Temperature Profile and Absorption Coefficient in One-Dimensional Semitransparent Medium by Inverse Radiation Analysis," *International Communications in Heat and Mass Transfer*, **27**, 5, pp. 635-643.

Liu, F. B., and Özisik, M. N, 1996, "Estimation of Inlet Temperature Profile in Laminar Duct Flow", *Inverse Problems in Engineering*, **3**, pp. 131-141.

Maruyama, S., 1993, "Uniform Isotropic Emission From an Involute Collector," *Journal of Heat Transfer*, **115**, 2, pp. 492-495.

Modest, M. F., 2003, *Radiative Heat Transfer*, 2<sup>nd</sup> Ed., Academic Press, New York, NY, pp. 644-653.

Morales, J. C., Harutunian, V., Oguma, M., and Howell, J. R., 1996, "Inverse Design of Radiating Enclosures with an Isothermal Participating Medium," in *Radiative Transfer I: Proc. of the First International Symposium on Radiation Transfer*, M. Pinar Menguç ed., Bengell House, New York, NY, pp. 579-593.

Morton, D. P., and Popova, E., 2001, "Monte Carlo Simulations for Stochastic Optimization," in *Encyclopedia of Optimization*, C. A. Floudas and P. M. Pardalos, eds., Kluwer Academic Publishers, Boston, MA, pp. 439-447.

Muschaweck, J., Spirkel, W., Timinger, A., Benz, N., Dörfler, M., Gut, M, and Kose, E., 2000, "Optimized Reflectors for Non-Tracking Solar Collectors with Tubular Absorbers," *Solar Energy*, **68**, 2, pp. 151-159.

Nash, S. G., and Sofer, A., 1996, *Linear and Nonlinear Programming*, McGraw Hill, New York, NY, pp. 396.



- Norman, S. A., 1992, "Optimization of Transient Temperature Uniformity in RTP Systems," *IEEE Transactions on Electron Devices*, **39**, 1, pp. 205-207.
- Nowak, I., Nowak, A. J., and Wrobel, L. C., 2002, "Identification of Phase Change Fronts by Bezier Splines and BEM," *International Journal of Thermal Sciences*, **41**, 6, pp. 492-499.
- Oguma, M., and Howell, J. R., 1995, "Solution of Two-Dimensional Blackbody Inverse Radiation Problems by Inverse Monte Carlo Method," *Proc. ASME/JSME Joint Thermal Engineering Conference*, Maui, HI.
- Özisik, M. N., and Orlande, H. R., B., 2000, *Inverse Heat Transfer: Fundamentals and Applications*, Taylor and Francis, New York, NY.
- Piegel L., and Tiller, W., 1997, *The NURBS Book*, 2<sup>nd</sup> Ed., Springer, Berlin.
- Pflug, G. C., 1996, *Optimization of Stochastic Models: The Interface Between Simulation and Optimization*, Kluwer Academic Press, Boston, MA, pp. 281-288.
- Ryan, J. P., Miller, J. A., and Burns, P. J., 1998, "Development and Use of a Monte Carlo Code to Conduct a Parametric Study of Cylindrical Solar Collector Arrays," *Proc. 1998 International Solar Energy Conference*, Albuquerque, NM, pp. 83-92.
- Siegel, R., and Howell, J. R., 2002a, *Thermal Radiation Heat Transfer*, 4<sup>th</sup> Ed., Taylor and Francis, New York, NY, pp. 390-406.
- Siegel, R., and Howell, J. R., 2002b, *Thermal Radiation Heat Transfer*, 4<sup>th</sup> Ed., Taylor and Francis, New York, NY, pp. 385-390.
- Sahinidis, N. V., 1996, "BARON: A General Purpose Global Optimization Software Package," *Journal of Global Optimization*, **8**, 2, pp. 201-205.
- Sasikumar, M., Balali, C., 2002, "A Holistic Optimization of Convecting-Radiating Fin Systems," *Journal of Heat Transfer*, **124**, 6, pp. 1110-1116.
- Tikhonov, A. N., 1975, "Inverse Problems in Heat Conduction," *Journal of Engineering Physics*, **29**, 1, pp. 816-820.
- Vogel, C. R., 2002, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, PA.

Ugray, Z., Lasdon, L., Plummer, J., Glover, F., Kelly, J., and Marti, R., 2002, "A Multistart Scatter Search Heuristic for Smooth NLP and MINLP Problems," *INFORMS Journal on Computing*, in press.

Winston, R., 1991, "Nonimaging Optics," *Scientific American*, **264**, 3, pp. 76-81.

Yoshitani, N., and Hasegawa, A., 1998, "Model-Based Control of Strip Temperature for the Heating Furnace in Continuous Annealing," *IEEE Transactions on Control Systems Technology*, **6**, 2, pp. 146-156.

## **Vita**

Kyle James Daun was born in Winnipeg, Canada, on December 11<sup>th</sup>, 1975, the son of Lynda and James Daun. He attended Grant Park High School from 1987 to 1993, and obtained his gliding and private pilot licenses through the Royal Canadian Air Cadets. In 1993 he was accepted into the mechanical engineering program at the University of Manitoba and received a degree of Bachelor of Science with distinction from that institution in 1997. In the same year, he entered the graduate program at the University of Waterloo and studied thermal radiation heat transfer under Professor Terry Hollands, which culminated in the degree of Masters in Applied Sciences in Mechanical Engineering in 1999.

He subsequently moved to Austin, Texas, to pursue a doctoral degree from the University of Texas at Austin under Professor Jack Howell. During this time, he was appointed twice as an assistant instructor of the undergraduate heat transfer course, served a term as the president of the Graduate Engineering Council, and received the NSERC PGS A Fellowship from the Natural Sciences and Engineering Council of Canada, and the Houston and Burton fellowships from the University of Texas at Austin.

Permanent address: 663 Beaverbrook St., Winnipeg, Manitoba,  
R3N 1N7, Canada

This dissertation was typed by the author.